Psychological Medicine

Longitudinal alterations in motivational salience processing in ultra high-risk subjects for psychosis

--Manuscript Draft--

Manuscript Number:	PSM-D-16-00500R2
Full Title:	Longitudinal alterations in motivational salience processing in ultra high-risk subjects for psychosis
Article Type:	Original Article
Corresponding Author:	André Schmidt, Ph.D. Institute of Psychiatry, Psychology and Neuroscience London, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Institute of Psychiatry, Psychology and Neuroscience
Corresponding Author's Secondary Institution:	
First Author:	André Schmidt, Ph.D.
First Author Secondary Information:	
Order of Authors:	André Schmidt, Ph.D.
	Mathilde Antoniades, M.Sc.
	Paul Allen, Professor
	Alice Egerton, Ph.D.
	Christopher C Chaddock, Ph.D.
	Stefan J Borgwardt, Professor
	Paolo Fusar-Poli, Professor
	Jonathan P Roiser, Professor
	Oliver Howes, Professor
	Philip McGuire, Professor
Order of Authors Secondary Information:	
Manuscript Region of Origin:	UNITED KINGDOM
Abstract:	Background: Impairments in the attribution of salience are thought to be fundamental to the development of psychotic symptoms and the onset of psychotic disorders. The aim of the present study was to explore longitudinal alterations in salience processing in ultra high-risk subjects for psychosis. Methods: 23 ultra high-risk subjects and 13 healthy controls underwent functional magnetic resonance imaging at two time points (mean interval of 17 months) while performing the Salience Attribution Test to assess neural responses to task-relevant (adaptive salience) and task-irrelevant (aberrant salience) stimulus features. Results: At presentation, high-risk subjects were less likely than controls to attribute salience to relevant features, and more likely to attribute salience to irrelevant stimulus features. These behavioural differences were no longer evident at follow-up. When attributing salience to relevant cue features, ultra high-risk subjects showed less activation than controls in the ventral striatum at both baseline and follow-up. Within the high-risk sample, amelioration of abnormal beliefs over the follow-up period was correlated with an increase in right ventral striatum activation during the attribution of salience to relevant cue features. Conclusions: These findings confirm that salience processing is perturbed in ultra high-risk subjects for psychosis, that this is linked to alterations in ventral striatum function,

	and that clinical outcomes are related to longitudinal changes in ventral striatum function during salience processing.

Longitudinal alterations in motivational salience processing in ultra high-risk subjects for psychosis

André Schmidt, Ph.D.[1], Mathilde Antoniades, M.Sc.[1], Paul Allen, Ph.D.[1,2], Alice Egerton, Ph.D.[1], Christopher A. Chaddock, Ph.D.[1], Stefan Borgwardt, M.D., PhD[1,3], Paolo Fusar-Poli, M.D., Ph.D.[1,4], Jonathan P. Roiser, Ph.D.[5], Oliver Howes, M.D., Ph.D.[1,6] and Philip McGuire, M.D., Ph.D.[1,4]

[1] Department of Psychosis Studies, King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK

[2] Department of Psychology, University of Roehampton, London, UK

[3] Department of Psychiatry (UPK), University of Basel, Switzerland

[4] OASIS Clinic, SLaM NHS Foundation Trust, London, UK

[5] Institute of Cognitive Neuroscience, University College London, UK

[6] Psychiatric Imaging, MRC Clinical Sciences Centre, Hammersmith Hospital, London, UK

Corresponding author: André Schmidt, Ph.D., Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, PO63 De Crespigny Park, London SE5 8AF, phone: +44 077 8666 6570, fax: +44 020 7848 0976, e-mail: andre.schmidt@kcl.ac.uk

Running title: Salience processing in UHR subjects

Word count abstract: 218

Word count text: 4462

Number of figures: 4

Number of tables: 1

Supplementary information: 1

Abstract

Background: Impairments in the attribution of salience are thought to be fundamental to the development of psychotic symptoms and the onset of psychotic disorders. The aim of the present study was to explore longitudinal alterations in salience processing in ultra high-risk subjects for psychosis.

Methods: 23 ultra high-risk subjects and 13 healthy controls underwent functional magnetic resonance imaging at two time points (mean interval of 17 months) while performing the Salience Attribution Test to assess neural responses to task-relevant (adaptive salience) and task-irrelevant (aberrant salience) stimulus features.

Results: At presentation, high-risk subjects were less likely than controls to attribute salience to relevant features, and more likely to attribute salience to irrelevant stimulus features. These behavioural differences were no longer evident at follow-up. When attributing salience to relevant cue features, ultra high-risk subjects showed less activation than controls in the ventral striatum at both baseline and follow-up. Within the high-risk sample, amelioration of abnormal beliefs over the follow-up period was correlated with an increase in right ventral striatum activation during the attribution of salience to relevant cue features.

Conclusions: These findings confirm that salience processing is perturbed in ultra high-risk subjects for psychosis, that this is linked to alterations in ventral striatum function, and that clinical outcomes are related to longitudinal changes in ventral striatum function during salience processing.

Declaration of interest

None.

Introduction

According to the aberrant salience model of psychosis (Heinz, 2002, Howes & Murray, 2014, Kapur, 2003), psychotic symptoms develop as a result of the inappropriate assignment of salience to contextually irrelevant internal and external experiences. This model is supported by evidence that patients with schizophrenia respond faster to task-irrelevant stimulus features than healthy controls (Pankow *et al.*, 2016), and that patients with prominent delusions rate irrelevant stimuli as more potentially rewarding than patients without delusions (Roiser *et al.*, 2009). This 'aberrant' attribution of salience is also evident in people at ultra high-risk (UHR) for psychosis, who are more likely to attribute salience to irrelevant stimulus features than healthy controls, with this tendency again related to the severity of abnormal beliefs (Roiser *et al.*, 2013).

Experiments in animals suggest that stimuli become motivationally salient when the release of dopamine in the striatum coincides with their perception (Kapur, 2003, Schultz *et al.*, 1997). In healthy individuals, aberrant salience measures are positively associated with ventral striatal (VS) presynaptic dopamine levels (Boehme *et al.*, 2015). Dopamine function in the striatum is abnormally elevated in both schizophrenia (Abi-Dargham *et al.*, 2000, Breier *et al.*, 1997, Howes *et al.*, 2012, Howes & Kapur, 2009, Kumakura *et al.*, 2007, Laruelle *et al.*, 1999, Laruelle *et al.*, 1996, Reith *et al.*, 1994) and UHR subjects (Egerton *et al.*, 2013, Howes *et al.*, 2011a, Howes *et al.*, 2011b, Howes *et al.*, 2009b, Mizrahi *et al.*, 2014) and the aberrant salience hypothesis proposes that this causes attribution of salience to irrelevant stimuli (Heinz & Schlagenhauf, 2010, Winton-Brown *et al.*, 2014). In addition, it has been hypothesised that because dopaminergic neurons may show more burst firing in psychosis (Goto & Grace, 2005, Winton-Brown *et al.*, 2014) the normal phasic dopaminergic response to relevant stimuli may become relatively diminished due to the high level of noise in the system (Heinz, 2002, Howes *et al.*, 2009a, Kapur, 2003). Psychosis may thus be associated with a reduced attribution of salience to relevant stimuli as well as increased attribution of salience to irrelevant stimuli. This is consistent with data from recent studies in UHR subjects and in patients with psychosis, which report impairments in both forms of salience processing (Pankow *et al.*, 2015, Roiser *et al.*, 2013, Roiser *et al.*, 2009).

Data from functional neuroimaging studies suggest that UHR subjects and patients with psychosis show altered activation in the VS during tasks that engage motivational salience processing. A recent meta-analysis suggested that reduced VS responses occur in patients with schizophrenia spectrum disorders relative to controls during the processing of contextually relevant information and that left VS hypoactivation was more severe in patients with high scores of negative symptoms (Radua *et al.*, 2015). The relation between VS activation during reward prediction and positive symptoms requires further investigation because only six studies were available (de Leeuw *et al.*, 2015, Esslinger *et al.*, 2012, Nielsen *et al.*, 2012a, Roiser *et al.*, 2013, Simon *et al.*, 2010, Wotruba *et al.*, 2014) and there was residual heterogeneity among them (Radua *et al.*, 2015). Interestingly, individual treatment with antipsychotics was associated with a normalization of VS activation during reward prediction, and this improvement was associated with the improvement of positive symptoms (Nielsen *et al.*, 2012b). With respect to contextually irrelevant information, it has been shown that striatal activation during incorrect distracter trials was positively correlated with aberrant salience symptoms in schizophrenia patients (Ceaser & Barch, 2015). In UHR subjects, the VS response to irrelevant stimulus features was found to be associated with the severity of abnormal beliefs (Roiser *et al.*, 2013). However, it is not know if altered VS activation during salience processing normalises in UHR individuals whose psychotic symptoms have remitted.

The Salience Attribution Test (SAT) is a paradigm that can be used to assess task-relevant and task-irrelevant motivational salience responses, termed adaptive and aberrant salience, respectively (Roiser *et al.*, 2009, Roiser *et al.*, 2010). Our objective was to assess the relationship between changes in clinical features in a UHR cohort and longitudinal changes in VS activation elicited during the SAT paradigm. Our first hypothesis was that at clinical presentation, UHR subjects would show increased aberrant but reduced adaptive salience processing compared to controls, and that these differences would be associated with concomitant alterations in VS activation. Our second hypothesis was that clinical improvements the UHR subjects subsequent to presentation would be associated with a longitudinal normalisation of behavioural and neural responses during salience processing.

Methods

Participants

Twenty-nine individuals who met the Comprehensive Assessment of At-Risk Mental States (CAARMS) (Yung *et al.*, 2005) criteria for the UHR state were recruited from Outreach and Support in South London (OASIS (Fusar-Poli *et al.*, 2013b)), a clinical service for people at high-risk for psychosis. According to international standard UHR criteria (see (Fusar-Poli *et al.*, 2013a) for a comprehensive review), inclusion required the presence of one or more of the following: (i) presence of attenuated psychosis symptoms (APS), (ii) genetic risk and deterioration syndrome (GRD), or (iii) brief limited and intermittent psychotic symptoms (BLIPS). Twenty-four individuals were included based on APS, three based on BLIPS and two based on APS+GRD. Following presentation, all subjects were provided with clinical care from OASIS (Fusar-Poli *et al.*, 2013b). Three subjects received antipsychotic medication and were thus excluded from the analysis. Twenty-six subjects also received cognitive behavioural therapy (CBT), which at the time of writing had been completed in nine subjects (twenty-six sessions in average, range 14-65). Seven subjects received low dose antidepressants: five citalopram (3 x 20mg, 1x 40 mg, and 1 x unknown dose), one mirtazapine (dose unknown) and one sertraline (100 mg). The specific treatments offered by OASIS have been detailed elsewhere (Fusar-Poli *et al.*, 2015). All subjects were managed in the community, attending regular outpatient appointments.

Fifteen healthy controls (HCs) from the same geographical area were recruited via local advertisements. Absence of psychiatric illness history was confirmed with the Mini International Neuropsychiatric Inventory (Sheehan *et al.*, 1998). None of the control subjects had a history of neurological illness, DSM-IV drug or alcohol dependence (Association., 2013). All subjects provided informed written consent to participate and the study was approved by the local National Health Service Research Ethics Committee.

The salience attribution task (SAT)

The SAT has been previously described in detail elsewhere (Roiser *et al.*, 2009, Roiser *et al.*, 2010). In brief, the SAT is a speeded-response game, rewarded with money, which measures responses to cue

features, which can be either task-relevant or task-irrelevant. On each trial of the task, participants were required to respond to a briefly presented square. Before the square appeared, a cue was shown indicating the likelihood of obtaining a reward for the forthcoming response. Participants received a monetary reward on 50% of trials, with more money awarded for faster responses. The cues varied in two different visual dimensions; colour (red or blue) and shape (animals or household objects). One of these cue dimensions was task-relevant and the other task-irrelevant. One task-relevant feature was highly associated with receiving a reward, with 87.5% of these trial types rewarded (e.g. *blue* stimuli). The other task-relevant dimension (e.g. *red* stimuli) was not rewarded on any trials. For the task-irrelevant dimension, an equal proportion of both features (e.g. *animal* and *household* stimuli) were rewarded. Participants were not informed about these contingencies, which remained the same over the two blocks of 64 trials within a testing session, and instead had to learn them over successive trials of the task. To avoid practice effects between baseline and follow-up, four different versions of the task were used, counterbalanced across participants, each with a different stimulus feature (blue, red, animal or household) rewarded with high probability.

Participants performed the task on two occasions, while being scanned using functional MRI. The baseline assessment was performed at the time of clinical presentation. The follow-up assessment was carried out approximately 17 months later. On each visit they performed the same version of the task twice. The SAT provides behavioural measures of adaptive (relevant) and aberrant (irrelevant) motivational salience on the basis of reaction times (RTs: implicit salience) and visual analogue scale ratings from 0% to 100% (VAS: explicit salience). Implicit adaptive salience is defined as the speeding of responses on high- relative to low-probability reward trials. Explicit adaptive salience is defined as the increase in VAS ratings on high- relative to low-probability reward trials. Implicit and explicit aberrant salience are defined as the absolute difference in RT and VAS rating, respectively, between the two levels of the task-irrelevant stimulus dimension (Roiser *et al.*, 2009).

Behavioural analysis

Behavioural scores on the SAT were analyzed using a repeated measures analysis of variance

(ANOVA) with time as within-subject and group as between-subject factors and years of education as a covariate. To test for group differences at baseline and follow-up separately, univariate ANOVA with education as covariate was used. Using box-and-whisker plots on each SAT measure for both groups separately, two HCs and two UHR subjects were excluded as outliers.

fMRI data acquisition and analysis

Scanning was performed on a whole-body 3T MRI General Electric (Milwaukee, Wisconsin) system. During each of the four scanning runs (two per day), we acquired T2*-weighted echo-planar images (EPIs) with the following parameters: 50 axial slices (sequential and top-down acquisition) of 2.4 mm thickness, 2.7 mm interslice gap, field of view 240 mm^2 and matrix size 64x64. The repetition time was 2.5 s and the echo time 25 ms. A total of 237 image volumes were acquired in a single functional run.

EPIs were analyzed using an event-related design with SPM12 (www.fil.ion.ucl.ac.uk/spm). Preprocessing was performed for each subject and time point separately. In brief, slice-timing correction was first performed on each volume using the middle slice as the reference. The images were then realigned to the first image in the series (following removal of dummy scans), spatially normalized to the Montreal Neurological Institute (MNI) template and smoothed with a Gaussian kernel of 8 mm full half-width maximum (FWHM). All images underwent visual inspection and participants with a high number of severely corrupted images and/or gross artefacts were excluded (two HCs and one UHRs). Additionally, all images were checked for movement artefacts, and all scans with more than 5 mm deviation from the previous scan in any dimension, resulting in corrupted volumes, were excluded and replaced with the average of the neighbouring volumes (5.1% in HCs and 1.5% in UHRs). Subjects with more than 10% corrupted volumes were excluded (two UHRs). In the final sample of 23 UHR subjects, 19 subjects were included based on BLIPS, three based on BLIPS and one based on APS+GRD.

Voxel-wise maximum likelihood parameter estimates were calculated during the first level analysis using the general linear model. Our design matrix included an autoregressive AR(1) model of serial correlations and a high-pass filter with a cutoff of 128 s. The onsets of each event were convolved

with the SPM synthetic hemodynamic response function. In this model, we included four 'cue' regressors, representing the different cue types and an 'outcome' regressor representing the time points when reward feedback was provided during the task. Cues on which participants failed to respond entirely were excluded from the analysis (regressor of no interest) due to the possibility that participants were not attending during the trial. Eight contrast images were generated per participant: adaptive and aberrant reward prediction at baseline and follow-up separately; average images over both visits for adaptive and aberrant reward prediction (to test for main effect of group); and two images subtracting the contrast vector of baseline from follow-up (to test for the main effect of time and time x group interaction). Adaptive reward prediction contrasts were defined as: high-probability reward cue features minus low-probability reward cue features across the task-relevant dimension. Aberrant reward prediction contrasts were defined as subjective "high-probability" reward cue features minus subjective "low-probability" reward cue features (based on the subject's VAS ratings for that run) across the task-irrelevant dimension (Roiser *et al.*, 2010).

Two-sample tests were conducted at the second level to test for group effects at baseline and follow-up separately, as well as to test for main effects of group and time and for time x group interactions. Significance was assessed at a cluster-level threshold of $p<0.05$ family-wise error (FWE) corrected across the whole brain, using an uncorrected cluster-forming threshold of $p<0.001$ (Petersson *et al.*, 1999, Woo *et al.*, 2014) with an extent threshold of 20 voxels. We also focused our analysis on the VS as this was part of our primary hypothesis, using a voxel-level approach. The VS region of interest was defined using coordinates taken from a previous SAT fMRI study in an independent UHR cohort (Roiser *et al.*, 2013): right ($x = 12$; $y = 12$; $z = –3$) and left ($x = –12$; $y = 9$; $z = –3$). Small volume correction was applied for this analysis using 15 mm spheres around these coordinates (Roiser *et al.*, 2013) and a voxel-level threshold of $p<0.05$ FWE corrected was considered significant. As groups differed in years of education, this variable was added as a covariate in the second-level model.

Relationships between brain activation, behaviour and symptoms

Relationships between neural responses and behavioural and clinical features were identified by

including outcome measures from the SAT, CAARMS and Global Assessment of Functioning (GAF) as covariates in second level models. The same procedure for correction for multiple comparisons as described above was employed. Relationships between behavioural salience responses and symptomatology in UHR subjects were tested with Pearson's correlation coefficients using Statistical Package for the Social Sciences (SPSS 16: SPSS Inc., Chicago, IL).

Results

Demographical and clinical features

The two groups did not differ in age, gender, handedness, IQ or cigarette, alcohol, cannabis and cocaine consumption, but HCs had more years of education (therefore, all group comparisons were co-varied for years of education). At baseline, UHR subjects had higher scores on CAARMS positive and negative symptoms and lower scores on the Global Assessment of Functioning (GAF). Over time, the UHR group showed significant improvements in CAARMS positive and negative symptoms, but not in GAF scores (Table 1).

Insert Table 1 about here

Behavioural data

Aberrant attribution of salience

Across both visits, UHR subjects showed significantly higher implicit aberrant salience than HCs subjects ($F(1,34)=6.718$, $p=0.014$), and there was a trend for a group x time interaction ($F(1,34)=3.225$, $p=0.081$). There was also a trend for a group x time interaction for explicit aberrant salience ($F(1,34)=3.325$, $p=0.077$). Based on our *a priori* hypotheses we constructed linear contrasts at each time point to test for the predicted group differences in aberrant salience.

At baseline, UHR subjects were more likely than HCs to attribute salience to irrelevant cue features (explicit aberrant salience) ($F(1,34)=4.732$, $p=0.037$), but did not exhibit greater implicit aberrant salience than HCs ($F(1,34)=0.964$, $p=0.333$). At follow-up the group difference in explicit aberrant salience was no longer significant ($F(1,34)=0.061$, $p=0.806$), but HCs had significantly lower implicit aberrant scores than the UHR group ($F(1,34)=12.296$, $p=0.001$) due to a reduction in this measure over time (Figure 1A and B).

Insert Figure 1 about here

Within the UHR group we detected no significant correlations between aberrant salience responses and psychotic symptoms (baseline, follow-up, change over time).

Adaptive attribution of salience

Across both visits, the UHR group had lower implicit adaptive salience scores than HCs ($F(1,34)=11.472$, $p=0.002$), as well as lower explicit adaptive salience scores ($F(1,34)=5.493$, $p=0.035$). There was also a significant group x time interaction for explicit adaptive salience ($F(1,34)=4.157$, $p=0.049$).

At baseline, UHR subjects had significantly lower implicit adaptive salience than HCs ($F(1,34)=13.866$, $p=0.001$) and also exhibited significantly lower explicit adaptive salience ($F(1,34)=9.043$, $p=0.005$). Both of these group differences were no longer significant at follow-up (implicit adaptive salience: $F(1,34)=3.733$, $p=0.062$; explicit adaptive salience: $F(1,34)=1.360$, $p=0.252$), due to improved scores in the UHR group together with relatively stable performance in HCs (Figure 2A and B).

Within the UHR group, explicit adaptive salience scores at follow-up were negatively correlated with the severity of abnormal beliefs ($r=-0.674$, $p<0.001$) (Supplementary Figure 1A) and of positive symptoms ($r=-0.653$, $p<0.001$) (Supplementary Figure 1B), and positively correlated with the level of global functioning ($r=0.497$, $p=0.014$) (Supplementary Figure 1C).

Insert Figure 2 about here

All behavioural results remained after excluding the UHR subject with a later transition to psychosis (Supplementary information 2A).

Activation during salience processing

Aberrant reward prediction

There were no significant effects of group or time, and no group x time interactions. There were also no significant group differences in responses to irrelevant cues at either baseline or follow-up.

Adaptive reward prediction

Across both time points, UHR subjects showed less activation than HCs in the VS, calcarine sulcus and midbrain bilaterally and in the left cuneus and middle temporal gyrus (main effect of group: Figure 3A, Supplementary Table 1). Across both groups, activation during adaptive reward prediction was greater at follow-up than at baseline in the bilateral VS and right thalamus (main effect of time: Figure 3B, Supplementary Table 2). No significant group x time interactions were found for adaptive reward prediction.

At baseline, the UHR group showed significantly less activation than HCs in the VS bilaterally and the left parahippocampal and middle temporal gyrus, and cerebellum during adaptive reward prediction (Supplementary Table 3). At follow-up, the UHR group continued to show significantly less activation in the VS bilaterally (Supplementary Table 4). All results remained after excluding the UHR subject with a later transition to psychosis (Supplementary information 2B).

Insert Figure 3 about here

There were no significant relationships between neural responses from the aberrant and adaptive reward prediction contrast and behavioural scores on the SAT (baseline, follow-up, change over time).

Relationship between longitudinal changes in clinical features and brain activation

Aberrant salience

There were no significant relationships between changes in clinical features and longitudinal changes in brain activation during aberrant reward prediction.

Adaptive salience

In the UHR group, there was a trend (t_{22}=1.775, p=0.09) for the mean severity of abnormal beliefs to improve between presentation and follow-up (Figure 4A). The degree of improvement in abnormal

beliefs over time was associated with the longitudinal increase in activation during adaptive reward prediction in the right VS and in the supplementary motor cortex bilaterally (Figure 4B and C, Supplementary Table 5). This relationship remained after excluding the UHR subject with a later transition to psychosis (Supplementary information 2B).

Insert Figure 4 about here

There were no significant correlations between longitudinal changes in negative symptoms and in neural responses during motivational salience processing.

Discussion

To our knowledge, this is the first longitudinal investigation of salience processing in subjects with psychotic symptoms. We explored the relationship between changes in the clinical features of people at UHR for psychosis after they had presented to clinical services and longitudinal changes in their behavioural and neural responses during aberrant and adaptive salience processing.

Aberrant salience

Consistent with the aberrant salience model (Heinz, 2002, Howes & Kapur, 2009, Kapur, 2003), we found that UHR subjects were more likely to attribute salience to irrelevant stimuli than HCs at clinical presentation. These data are consistent with a previous report of increased explicit aberrant salience in an independent UHR sample (Roiser *et al.*, 2013). A study using the SAT in first-episode schizophrenia did not find a difference in the patient sample overall, but found that aberrant salience was related to the severity of delusions and negative symptoms within the patient group (Roiser *et al.*, 2009). However, it should be noted that another study found no significant differences between UHR subjects, first-episode patients and controls in aberrant salience attribution (Smieskova *et al.*, 2015).

During the 17-month follow-up period, there was a reduction in explicit aberrant salience in UHR subjects, such that there was no longer a significant group difference relative to controls. On the basis that abnormal salience processing is proposed to underlie the generation of psychotic symptoms (Roiser *et al.*, 2013, Roiser *et al.*, 2009), we tested whether longitudinal changes in aberrant salience processing were related to changes in clinical features in the UHR subjects during the follow-up period. Although UHR subjects showed improvements clinically, there were no significant correlations between changes in these variables and longitudinal changes in behavioural measures of aberrant salience processing. It has been proposed that the link between aberrant salience and symptoms is moderated by cognitive biases (Howes & Murray, 2014), which may account for the lack of direct relationship between aberrant salience and symptom change in our data.

Adaptive salience

The aberrant salience model proposed that adaptive salience is intact in patients with psychosis, but may become impaired as a result of treatment with antipsychotic medication (Heinz, 2002, Kapur, 2003). The first experimental study of salience processing in first-episode psychosis using the SAT found that patients showed impaired adaptive salience, and this was attributed to be an effect of antipsychotic treatment (Roiser *et al.*, 2009). However, a subsequent study of largely medication-naïve UHR subjects also found a trend for reduced implicit adaptive salience (Roiser *et al.*, 2013). In the present study, which involved a larger patient sample, at presentation, UHR subjects showed significantly reduced adaptive salience responses. As all of our UHR subjects were naïve to antipsychotic medication at this stage, these data not only suggest that adaptive salience is impaired in UHR subjects, but that this is not secondary to antipsychotic treatment. Consistent with this interpretation, a recent study found that adaptive salience processing was numerically (though not significantly) impaired in first-episode psychosis patients, but this impairment was if anything less marked in antipsychotic-treated than untreated patients (Smieskova *et al.*, 2015).

Although significant behavioural differences in adaptive salience processing were only present at baseline, group differences in activation during adaptive salience processing were seen at both presentation and follow-up time-points. At both time points, UHR subjects showed reduced activation relative to controls in the VS. This is consistent with a recent meta-analysis demonstrating reduced VS activity in response to reward-predicting cues in schizophrenia spectrum disorders (Radua *et al.*, 2015), and reports of altered VS activation in patients with psychosis during reward prediction error tasks (Gradin *et al.*, 2011, Murray *et al.*, 2008). Furthermore, within the UHR group, improvement in abnormal beliefs over the follow-up period was correlated with the degree to which VS activation increased over time during adaptive reward prediction. This finding is in line with data from unmedicated first-episode patients demonstrating a negative correlation between the severity of delusional symptoms and reward prediction signals in the VS (Esslinger *et al.*, 2012). Taken with longitudinal PET imaging findings that changes in dopamine synthesis capacity in dorsal (associative) striatum are associated with change in clinical state (Howes *et al.*, 2011a), our findings suggest that alterations in both ventral and dorsal striatum are linked to symptom change. A possible mechanism

could be that hyperactive inputs from the hippocampus to the ventral striatum in psychosis may impact dopaminergic neurons that project to more dorsal (associative) striatal areas and thereby affect dorsal striatum related salience processing (Haber, 2003, Lodge & Grace, 2011, 2012, Modinos *et al.*, 2015). The amelioration of abnormal beliefs in UHR subjects was also associated with longitudinal increases in activation in the supplementary motor cortex to reward predicting cue features. The latter finding was not predicted, as supplementary motor cortex is not specifically implicated in motivational salience processing. However, the SAT is a complex task that also involves sustained attention, maintaining stimulus information in memory, decision-making and response selection (Roiser *et al.*, 2013), and the UHR state is associated with a broad range of cognitive impairments (Fusar-Poli *et al.*, 2012). We therefore speculate that this finding in the supplementary motor cortex may be related to alterations in one or more of these processes, possibly secondary to changes in striatal function. Furthermore, UHR subjects also showed reduced activation in the calcarine sulcus, cuneus, midbrain and middle temporal gyrus across both visits during the attribution of salience to relevant stimuli, as well as reduced activation in the parahippocampal gyrus, cerebellum, midbrain, middle temporal gyrus, middle and anterior cingulate cortex, inferior frontal gyrus and insula at baseline and/or follow-up (see supplementary information 1 and 2B for more details). Together with the striatum, integration of these regions is important to sustain emotion and cognition, especially during the detection and processing of salient information (Menon & Uddin, 2010, Seeley *et al.*, 2007). Dysfunction of this network and abnormal network switching when dealing with a relevant task at hand has been proposed to underlie the formation of psychotic symptoms (Palaniyappan & Liddle, 2012, Palaniyappan *et al.*, 2013, Schmidt *et al.*, 2016).

Some limitations of our study merit comment. The sample sizes were modest, largely because inclusion required that participants completed multi-modal neuroimaging assessments at both baseline and follow-up. The modest group sizes may thus have accounted for the absence of significant group differences in activation during aberrant salience processing. A further consideration is that at the time of writing, only one UHR subject had developed a psychotic disorder (all results remained after excluding this subject, see supplementary information 2A and B for details), precluding any

examination of the relationship between abnormal salience processing and the risk of transition to psychosis. In this regard, it is possible that the low conversion rate in our UHR sample may explain the lack of alterations in brain activation during aberrant salience processing. Future large-scale studies with a meaningful ratio between converters and non-converters are required to test if functional brain alterations during aberrant reward prediction are evident in UHR subjects who later develop psychosis or if the risk of transition to psychosis is more related to impaired activation when dealing with a relevant task at hand (i.e. adaptive reward prediction). Furthermore, in accordance with aberrant salience model (Kapur, 2003), the SAT has been designed to measure abnormal motivational (reward) salience processing in psychosis and its relation to dopamine dysregulation in the VS. However, motivation is not the only form of salience (Winton-Brown *et al.*, 2014), and it would be important to test ventral and dorsal (associative) striatal activation in psychosis during other forms of salience processing that are not measured using speeded response tasks. Finally, subsequent to presentation, some of the UHR subjects received CBT or low doses of antidepressants, which may have influenced our findings. In this study, the numbers of subjects receiving different forms of treatment were too small to allow for meaningful sub-group analyses and this issue would be better addressed in longitudinal studies that were explicitly linked to a clinical trial of an intervention that might be expected to improve motivational salience processing.

In summary, this study shows that UHR subjects exhibit behavioural deficits in both adaptive and aberrant salience processing at clinical presentation, which disappeared along with the remission of attenuated psychotic symptoms over the follow-up period. Our results further indicate ventral striatal hypoactivation in UHR subjects during adaptive reward prediction at baseline and follow-up and that the amelioration of abnormal beliefs over the follow-up period is linked to a longitudinal increase in VS activation during adaptive reward prediction.

Acknowledgements

This work was supported by the Swiss National Science Foundation (No. P2ZHP3_155184 to Dr Schmidt), the Medical Research Council (MRC) UK research grant (No. G0700995 to Drs Allen, Howes and McGuire). Dr Howes has received investigator-initiated research funding from and/or participated in advisory/ speaker meetings organised by Astra-Zeneca, Autifony, BMS, Eli Lilly, Heptares, Jansenn, Lundbeck, Lyden-Delta, Otsuka, Servier, Sunovion, Rand and Roche.

Note

Supplementary material accompanies this paper on the Journal's website (http://journals.cambridge.org).

Declaration of Interest

None.

References

Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L. S., Weiss, R., Cooper, T. B., Mann, J. J., Van Heertum, R. L., Gorman, J. M. & Laruelle, M. (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proceedings of the National Academy of Sciences USA* **97**, 8104-9.

APA (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5.* American Psychiatric Association: Washington.

Boehme, R., Deserno, L., Gleich, T., Katthagen, T., Pankow, A., Behr, J., Buchert, R., Roiser, J. P., Heinz, A. & Schlagenhauf, F. (2015). Aberrant Salience Is Related to Reduced Reinforcement Learning Signals and Elevated Dopamine Synthesis Capacity in Healthy Adults. *Journal of Neuroscience* **35**, 10103-11.

Breier, A., Su, T. P., Saunders, R., Carson, R. E., Kolachana, B. S., de Bartolomeis, A., Weinberger, D. R., Weisenfeld, N., Malhotra, A. K., Eckelman, W. C. & Pickar, D. (1997). Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: evidence from a novel positron emission tomography method. *Proceedings of the National Academy of Sciences USA* **94**, 2569-74.

Ceaser, A. E. & Barch, D. M. (2015). Striatal Activity is Associated with Deficits of Cognitive Control and Aberrant Salience for Patients with Schizophrenia. *Frontiers in Human Neuroscience* **9**, 687.

de Leeuw, M., Kahn, R. S. & Vink, M. (2015). Fronto-striatal dysfunction during reward processing in unaffected siblings of schizophrenia patients. *Schizophrenia Bulletin* **41**, 94-103.

Egerton, A., Chaddock, C. A., Winton-Brown, T. T., Bloomfield, M. A., Bhattacharyya, S., Allen, P., McGuire, P. K. & Howes, O. D. (2013). Presynaptic striatal dopamine dysfunction in people at ultra-high risk for psychosis: findings in a second cohort. *Biological Psychiatry* **74**, 106-12.

Esslinger, C., Englisch, S., Inta, D., Rausch, F., Schirmbeck, F., Mier, D., Kirsch, P., Meyer-Lindenberg, A. & Zink, M. (2012). Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophrenia Research* **140**, 114-21.

Fusar-Poli, P., Borgwardt, S., Bechdolf, A., Addington, J., Riecher-Rossler, A., Schultze-Lutter, F., Keshavan, M., Wood, S., Ruhrmann, S., Seidman, L. J., Valmaggia, L., Cannon, T., Velthorst, E., De Haan, L., Cornblatt, B., Bonoldi, I., Birchwood, M., McGlashan, T., Carpenter, W., McGorry, P., Klosterkotter, J., McGuire, P. & Yung, A. (2013a). The Psychosis High-Risk State A Comprehensive State-of-the-Art Review. *JAMA Psychiatry* **70**, 107-120.

Fusar-Poli, P., Byrne, M., Badger, S., Valmaggia, L. R. & McGuire, P. K. (2013b). Outreach and support in south London (OASIS), 2001-2011: ten years of early diagnosis and treatment for young individuals at high clinical risk for psychosis. *European Psychiatry* **28**, 315-26.

Fusar-Poli, P., Deste, G., Smieskova, R., Barlati, S., Yung, A. R., Howes, O., Stieglitz, R. D., Vita, A., McGuire, P. & Borgwardt, S. (2012). Cognitive Functioning in Prodromal Psychosis: A Meta-analysisCognitive Functioning in Prodromal Psychosis. *Archives of General Psychiatry* **69**, 562-71.

Fusar-Poli, P., Frascarelli, M., Valmaggia, L., Byrne, M., Stahl, D., Rocchetti, M., Codjoe, L., Weinberg, L., Tognin, S., Xenaki, L. & McGuire, P. (2015). Antidepressant, antipsychotic and psychological interventions in subjects at high clinical risk for psychosis: OASIS 6-year naturalistic study. *Psychological Medicine* **45**, 1327-39.

Goto, Y. & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience* **8**, 805-12.
Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J. & Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* **134**, 1751-64.
Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy* **26**, 317-30.
Heinz, A. (2002). Dopaminergic dysfunction in alcoholism and schizophrenia--psychopathological and behavioral correlates. *European Psychiatry* **17**, 9-16.
Heinz, A. & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophrenia Bulletin* **36**, 472-85.
Howes, O., Bose, S., Turkheimer, F., Valli, I., Egerton, A., Stahl, D., Valmaggia, L., Allen, P., Murray, R. & McGuire, P. (2011a). Progressive increase in striatal dopamine synthesis capacity as patients develop psychosis: a PET study. *Molecular Psychiatry* **16**, 885-6.
Howes, O. D., Bose, S. K., Turkheimer, F., Valli, I., Egerton, A., Valmaggia, L. R., Murray, R. M. & McGuire, P. (2011b). Dopamine synthesis capacity before onset of psychosis: a prospective [18F]-DOPA PET imaging study. *American Journal of Psychiatry* **168**, 1311-7.
Howes, O. D., Egerton, A., Allan, V., McGuire, P., Stokes, P. & Kapur, S. (2009a). Mechanisms underlying psychosis and antipsychotic treatment response in schizophrenia: insights from PET and SPECT imaging. *Current Pharmaceutical Design* **15**, 2550-9.
Howes, O. D., Kambeitz, J., Kim, E., Stahl, D., Slifstein, M., Abi-Dargham, A. & Kapur, S. (2012). The nature of dopamine dysfunction in schizophrenia and what this means for treatment. *Archives of General Psychiatry* **69**, 776-86.
Howes, O. D. & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version III--the final common pathway. *Schizophrenia Bulletin* **35**, 549-62.
Howes, O. D., Montgomery, A. J., Asselin, M. C., Murray, R. M., Valli, I., Tabraham, P., Bramon-Bosch, E., Valmaggia, L., Johns, L., Broome, M., McGuire, P. K. & Grasby, P. M. (2009b). Elevated striatal dopamine function linked to prodromal signs of schizophrenia. *Archives of General Psychiatry* **66**, 13-20.
Howes, O. D. & Murray, R. M. (2014). Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet* **383**, 1677-87.
Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry* **160**, 13-23.
Kumakura, Y., Cumming, P., Vernaleken, I., Buchholz, H. G., Siessmeier, T., Heinz, A., Kienast, T., Bartenstein, P. & Gründer, G. (2007). Elevated [18F]fluorodopamine turnover in brain of patients with schizophrenia: an [18F]fluorodopa/positron emission tomography study. *Journal of Neuroscience* **27**, 8080-7.
Laruelle, M., Abi-Dargham, A., Gil, R., Kegeles, L. & Innis, R. (1999). Increased dopamine transmission in schizophrenia: relationship to illness phases. *Biological Psychiatry* **46**, 56-72.
Laruelle, M., Abi-Dargham, A., van Dyck, C. H., Gil, R., D'Souza, C. D., Erdos, J., McCance, E., Rosenblatt, W., Fingado, C., Zoghbi, S. S., Baldwin, R. M., Seibyl, J. P., Krystal, J. H., Charney, D. S. & Innis, R. B. (1996). Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proceedings of the National Academy of Sciences USA* **93**, 9235-40.

Lodge, D. J. & Grace, A. A. (2011). Hippocampal dysregulation of dopamine system function and the pathophysiology of schizophrenia. *Trends in Pharmacological Sciences* **32**, 507-13.
Lodge, D. J. & Grace, A. A. (2012). Divergent activation of ventromedial and ventrolateral dopamine systems in animal models of amphetamine sensitization and schizophrenia. *International Journal of Neuropsychopharmacology* **15**, 69-76.
Menon, V. & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function* **214**, 655-67.
Mizrahi, R., Kenk, M., Suridjan, I., Boileau, I., George, T. P., McKenzie, K., Wilson, A. A., Houle, S. & Rusjan, P. (2014). Stress-induced dopamine response in subjects at clinical high risk for schizophrenia with and without concurrent cannabis use. *Neuropsychopharmacology* **39**, 1479-89.
Modinos, G., Allen, P., Grace, A. A. & McGuire, P. (2015). Translating the MAM model of psychosis to humans. *Trends in Neurosciences* **38**, 129-38.
Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., Jones, P. B., Bullmore, E. T., Robbins, T. W. & Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry* **13**, 239, 267-76.
Nielsen, M., Rostrup, E., Wulff, S., Bak, N., Lublin, H., Kapur, S. & Glenthøj, B. (2012a). Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biological Psychiatry* **71**, 898-905.
Nielsen, M. O., Rostrup, E., Wulff, S., Bak, N., Broberg, B. V., Lublin, H., Kapur, S. & Glenthoj, B. (2012b). Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Archives of General Psychiatry* **69**, 1195-204.
Palaniyappan, L. & Liddle, P. F. (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *Journal of Psychiatry & Neuroscience* **37**, 17-27.
Palaniyappan, L., Simmonite, M., White, T. P., Liddle, E. B. & Liddle, P. F. (2013). Neural primacy of the salience processing system in schizophrenia. *Neuron* **79**, 814-28.
Pankow, A., Katthagen, T., Diner, S., Deserno, L., Boehme, R., Kathmann, N., Gleich, T., Gaebler, M., Walter, H., Heinz, A. & Schlagenhauf, F. (2016). Aberrant Salience Is Related to Dysfunctional Self-Referential Processing in Psychosis. *Schizophrenia Bulletin* **42**, 67-76.
Petersson, K. M., Nichols, T. E., Poline, J. B. & Holmes, A. P. (1999). Statistical limitations in functional neuroimaging. II. Signal detection and statistical inference. *Philosophical Transactions of the Royal Society B: Biological sciences* **354**, 1261-81.
Radua, J., Schmidt, A., Borgwardt, S., Heinz, A., Schlagenhauf, F., McGuire, P. & Fusar-Poli, P. (2015). Ventral Striatal Activation During Reward Processing in Psychosis: A Neurofunctional Meta-Analysis. *JAMA Psychiatry* **72**, 1243-51.
Reith, J., Benkelfat, C., Sherwin, A., Yasuhara, Y., Kuwabara, H., Andermann, F., Bachneff, S., Cumming, P., Diksic, M., Dyve, S. E., Etienne, P., Evans, A. C., Lal, S., Shevell, M., Savard, G., Wong, D. F., Chouinard, G. & Gjedde, A. (1994). Elevated dopa decarboxylase activity in living brain of patients with psychosis. *Proceedings of the National Academy of Sciences USA* **91**, 11651-4.
Roiser, J. P., Howes, O. D., Chaddock, C. A., Joyce, E. M. & McGuire, P. (2013). Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophrenia Bulletin* **39**, 1328-36.
Roiser, J. P., Stephan, K. E., den Ouden, H. E., Barnes, T. R., Friston, K. J. & Joyce, E. M. (2009). Do patients with schizophrenia exhibit aberrant salience? *Psychological Medicine* **39**, 199-209.

Roiser, J. P., Stephan, K. E., den Ouden, H. E., Friston, K. J. & Joyce, E. M. (2010). Adaptive and aberrant reward prediction signals in the human brain. *Neuroimage* **50**, 657-64.
Schmidt, A., Palaniyappan, L., Smieskova, R., Simon, A., Riecher-Rössler, A., Lang, U. E., Fusar-Poli, P., McGuire, P. & Borgwardt, S. J. (2016). Dysfunctional insular connectivity during reward prediction in patients with first-episode psychosis. *Journal of Psychiatry & Neuroscience* **41**, 150234.
Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* **275**, 1593-9.
Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L. & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience* **27**, 2349-56.
Sheehan, D., Lecrubier, Y., Sheehan, K., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R. & Dunbar, G. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59**, 22-33.
Simon, J. J., Biller, A., Walther, S., Roesch-Ely, D., Stippich, C., Weisbrod, M. & Kaiser, S. (2010). Neural correlates of reward processing in schizophrenia--relationship to apathy and depression. *Schizophrenia Research* **118**, 154-61.
Smieskova, R., Roiser, J. P., Chaddock, C. A., Schmidt, A., Harrisberger, F., Bendfeldt, K., Simon, A., Walter, A., Fusar-Poli, P., McGuire, P. K., Lang, U. E., Riecher-Rössler, A. & Borgwardt, S. (2015). Modulation of motivational salience processing during the early stages of psychosis. *Schizophrenia Research* **166,** 17-23.
Winton-Brown, T. T., Fusar-Poli, P., Ungless, M. A. & Howes, O. D. (2014). Dopaminergic basis of salience dysregulation in psychosis. *Trends in Neurosciences* **37**, 85-94.
Woo, C. W., Krishnan, A. & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* **91**, 412-9.
Wotruba, D., Heekeren, K., Michels, L., Buechler, R., Simon, J. J., Theodoridou, A., Kollias, S., Rössler, W. & Kaiser, S. (2014). Symptom dimensions are associated with reward processing in unmedicated persons at risk for psychosis. *Frontiers in Behavioral Neuroscience* **8**, 382.
Yung, A. R., Yuen, H. P., McGorry, P. D., Phillips, L. J., Kelly, D., Dell'Olio, M., Francey, S. M., Cosgrave, E. M., Killackey, E., Stanford, C., Godfrey, K. & Buckby, J. (2005). Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *Australian & New Zealand Journal of Psychiatry* **39**, 964-71.

Tables

Table 1. Demographical and clincal characteristics of the study sample.

	Healthy controls (n=13)	Ultra high-risk subjects (n=23)	Group statistics[+*]
At Baseline			
Age (years, mean ± SD, [range])	24.38 ± 5.32 [20-36]	21.57 ± 3.55 [18-29]	t_{34}=1.709, p=0.104[+]
Female/male (n)	3/10	11/12	χ^2=2.141, p=0.143[+]
Handedness (R/L)	11/2	22/1	χ^2=1.324, p=0.250[+]
Education (years, mean ± SD)	14.77 ± 1.64	12.52 ± 2.25	t_{34}=3.436, **p=0.002**[+]
NART FSIQ (mean ± SD)	111.50 ± 6.10	107.04 ± 9.07	t_{34}=1.578, p=0.124[+]
Cigarettes per day (mean ± SD)	2.69 ± 4.38	5.57 ± 8.13	t_{34}=1.376, p=0.178[+]
Alcohol units per week (mean ± SD)	11 ± 10.52	6.87 ± 9.28	t_{34}=1.179, p=0.251[+]
Cannabis consumption (yes/no)	8/5	14/9	χ^2=0.002, p=0.0968[+]
Cocaine consumption (yes/no)	4/9	8/15	χ^2=0.060, p=0.806[+]
GAF (mean ± SD)	84.15 ± 4.88	59.74 ± 7.61	t_{34}=11.706, **p<0.001**[+]
CAARMS Positive symptoms (mean ± SD)	0.54 ± 1.20	7.65 ± 3.81	t_{34}=-8.262, **p<0.001**[+]
CAARMS Negative symptoms (mean ± SD)	0.17 ± 0.58	6.39 ± 3.29	t_{34}=-7.025, **p<0.001**[+]
At Follow-up			
Age (years, mean ± SD, [range])	25.70 ± 5.33 [21-37]	22.96 ± 3.48 [19-30]	t_{34}=1.668, p=0.113[+]
Cigarettes per day (mean ± SD)	/	3.13 ± 5.65	t_{22}=1.611, p=0.121[*]
Alcohol units per week (mean ± SD)	/	7.74 ± 10.82	t_{22}=0.506, p=0.618[*]
Cannabis consumption (yes/no)	/	14/9	χ^2=0.000, p=1[*]
Cocaine consumption (yes/no)	/	8/15	χ^2=0.000, p=1[*]
GAF (mean ± SD)	/	62.39 ± 15.78	t_{22}=0.896, p=0.38[*]
CAARMS Positive symptoms (mean ± SD)	/	5.22 ± 4.88	t_{22}=2.811, **p=0.010**[*]
CAARMS Negative symptoms (mean ± SD)	/	4.22 ± 4.35	t_{17}=2.663, **p=0.016**[*]

CAARMS, Comprehensive Assessment of At-Risk Mental States. CAARMS-positive symptoms were the sum of severity scores for unusual thought content

(abnormal belief), non-bizarre ideas, perceptual abnormalities and disorganized speed; negative symptoms were the sum of severity scores for alogia, avolition/apathy and anhedonia. GAF, Global Assessment of Functioning; NART, National Adult Reading Test Full Scale IQ; SD, standard deviation.

[+] 2-sample t tests and chi-squared tests between groups, respectively.

[*] paired tests and chi-squared tests within ultra high-risk subjects between baseline and follow-up assessment.

Figure legends

Figure 1. (A) *Implicit* (reaction times, msec) and **(B)** *explicit* (visual analogue scale) scores for *aberrant* motivational salience processing in healthy controls and subjects at ultra high-risk (UHR) for psychosis.

Figure 2. (A) *Implicit* (reaction times, msec) and **(B)** *explicit* (visual analogue scale) scores for *adaptive* motivational salience processing in healthy controls and subjects at ultra high-risk (UHR) for psychosis.

Figure 3. (A) Greater activation during adaptive reward prediction in healthy controls (HC) compared with ultra high-risk (UHR) subjects across both visits. **(B)** Greater activation at follow-up relative to baseline during adaptive reward prediction across both groups. Images are displayed at a cluster-forming threshold of $p<0.001$ uncorrected, with an extent threshold of 20 voxels. Colour bars indicate t-values.

Figure 4. (A) Unusual Thought Content (abnormal beliefs) at baseline (mean: 3.04) and follow-up (mean: 2.09) in ultra high-risk (UHR) subjects ($t_{22}=1.775$, $p=0.09$). **(B)** Negative correlation between changes in brain activation during adaptive reward prediction and changes in abnormal beliefs from baseline to follow-up in UHR subjects. The image is displayed at a cluster-forming threshold of $p<0.001$ uncorrected, with an extent threshold of 20 voxels. The colour bar indicates t-values. **(C)** Scatter plot of negative relationship between change in right ventral striatum activation during adaptive salience processing (taken from the peak voxel in B) and change in abnormal beliefs (CAARMS item unusual thought content) from baseline to follow-up in UHR subjects ($r=-0.702$).

Figure1A

Click here to download Figure(s) Figure1A.tiff

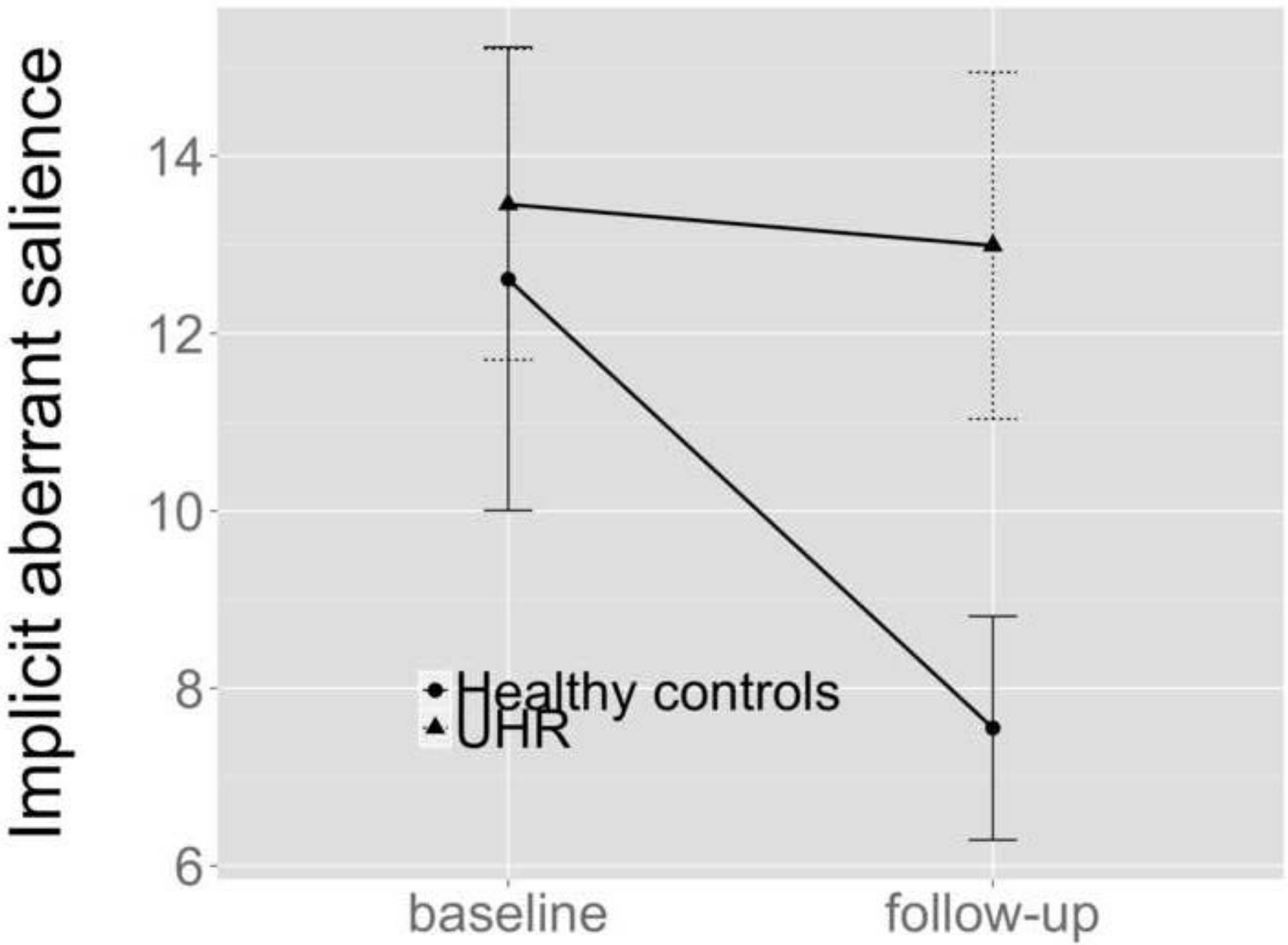


Figure1B

Click here to download Figure(s) Figure1B.tiff

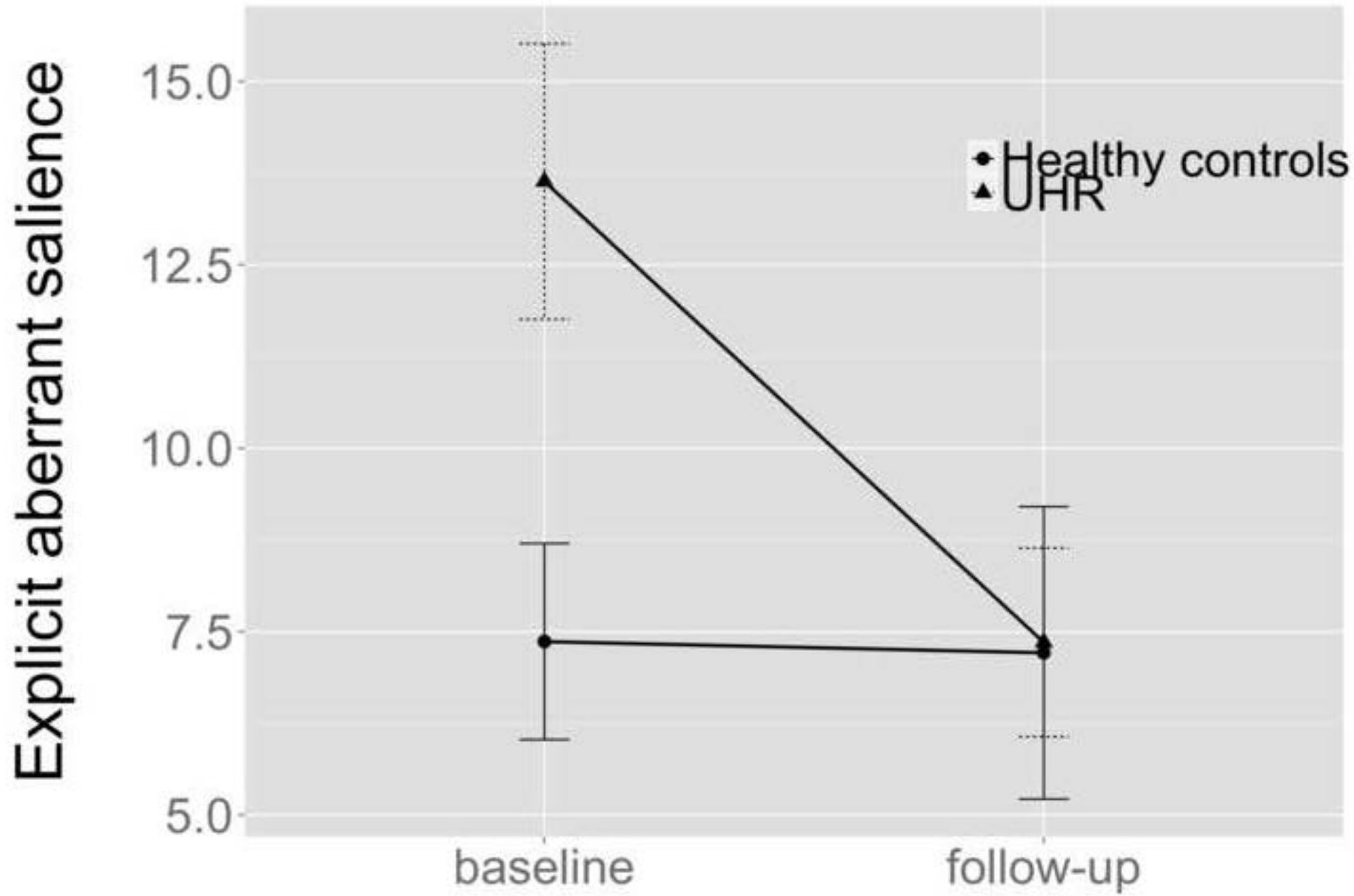


Figure2A

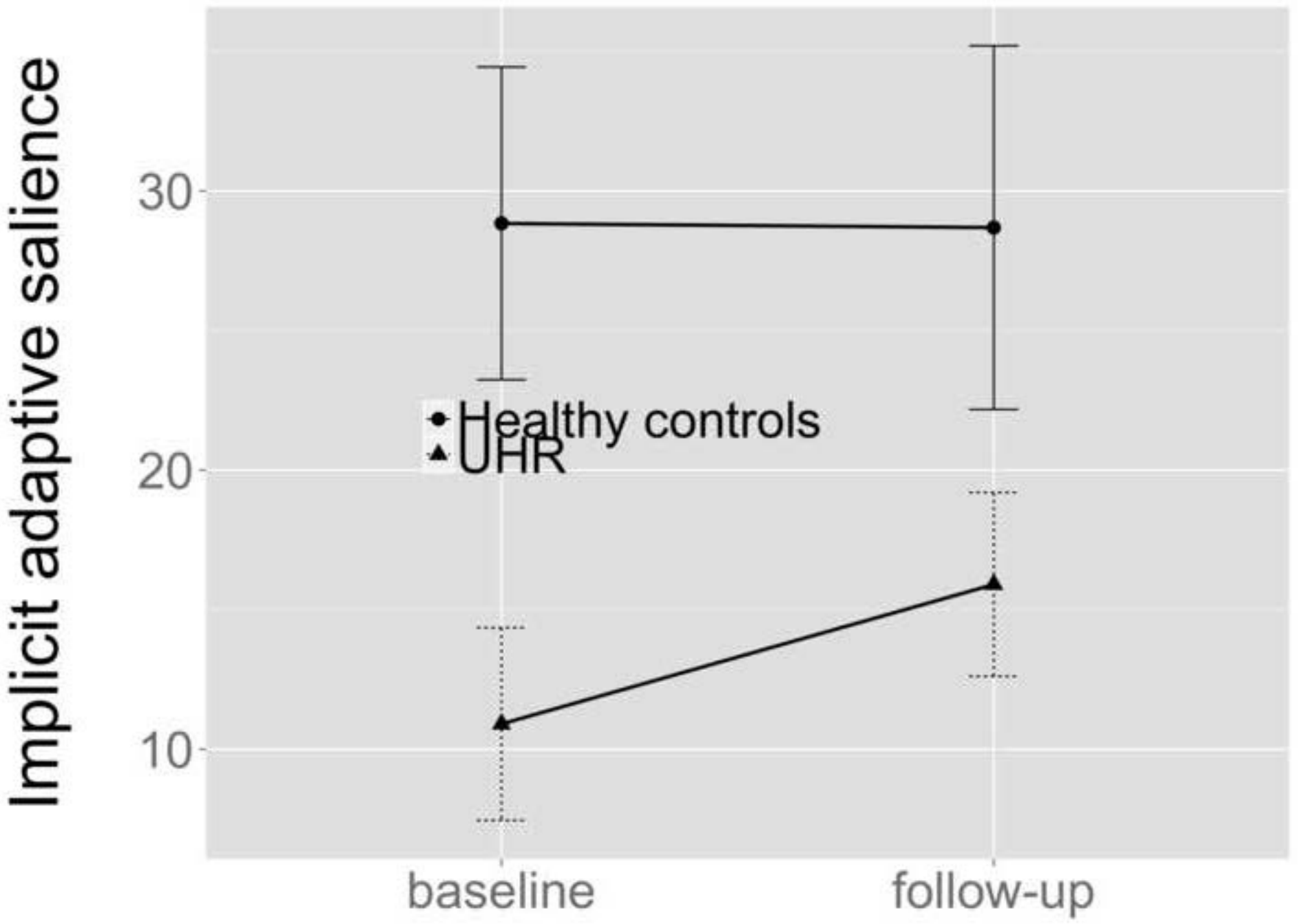
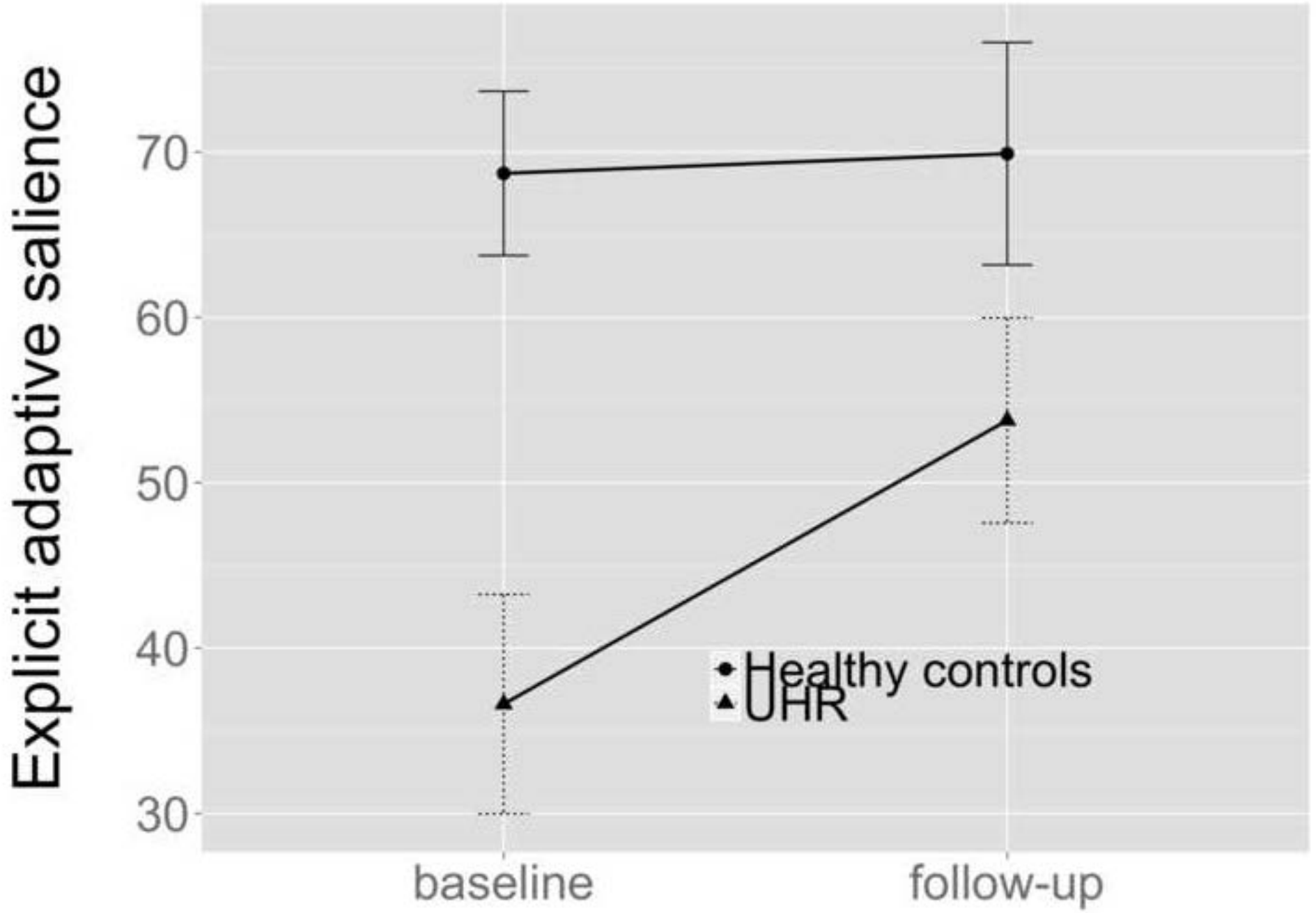
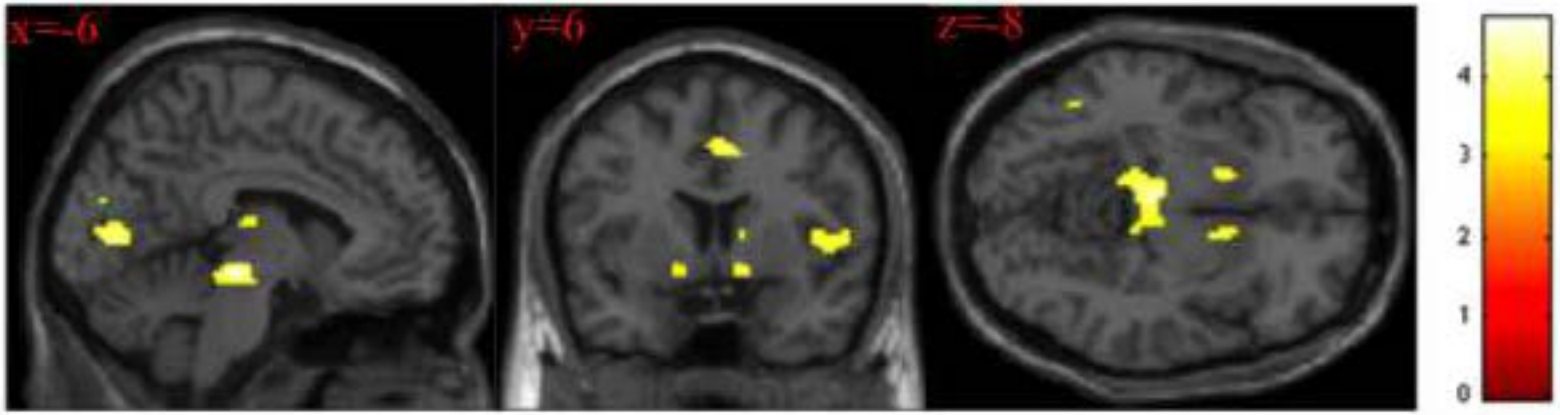


Figure2B



A) **Group effect (HC > UHR) during adaptive reward prediction**



B) **Time effect (follow-up > baseline) during adaptive reward prediction**

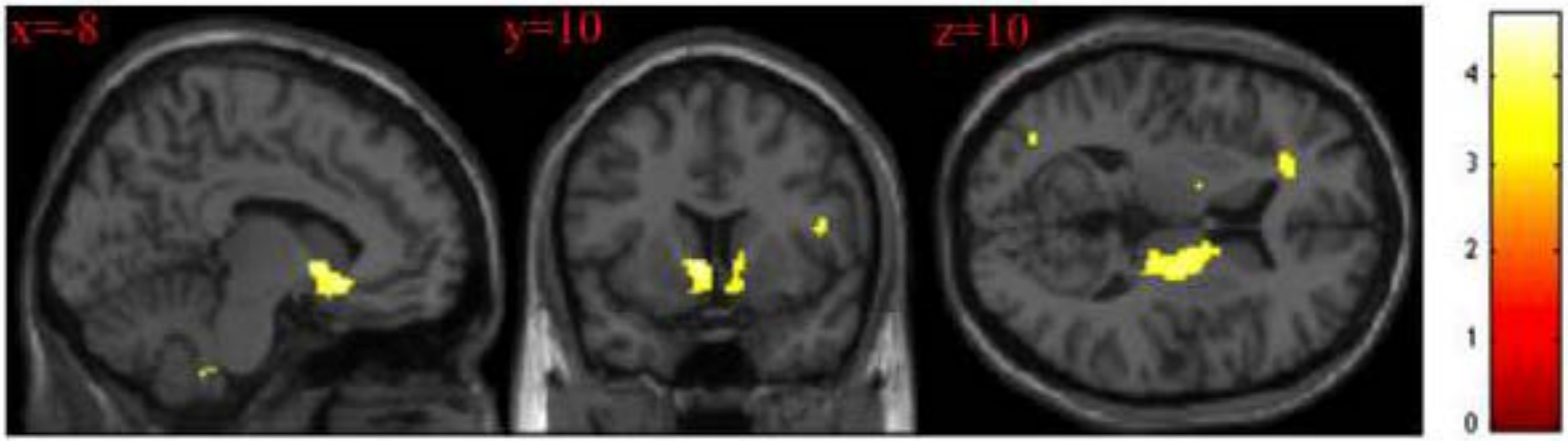
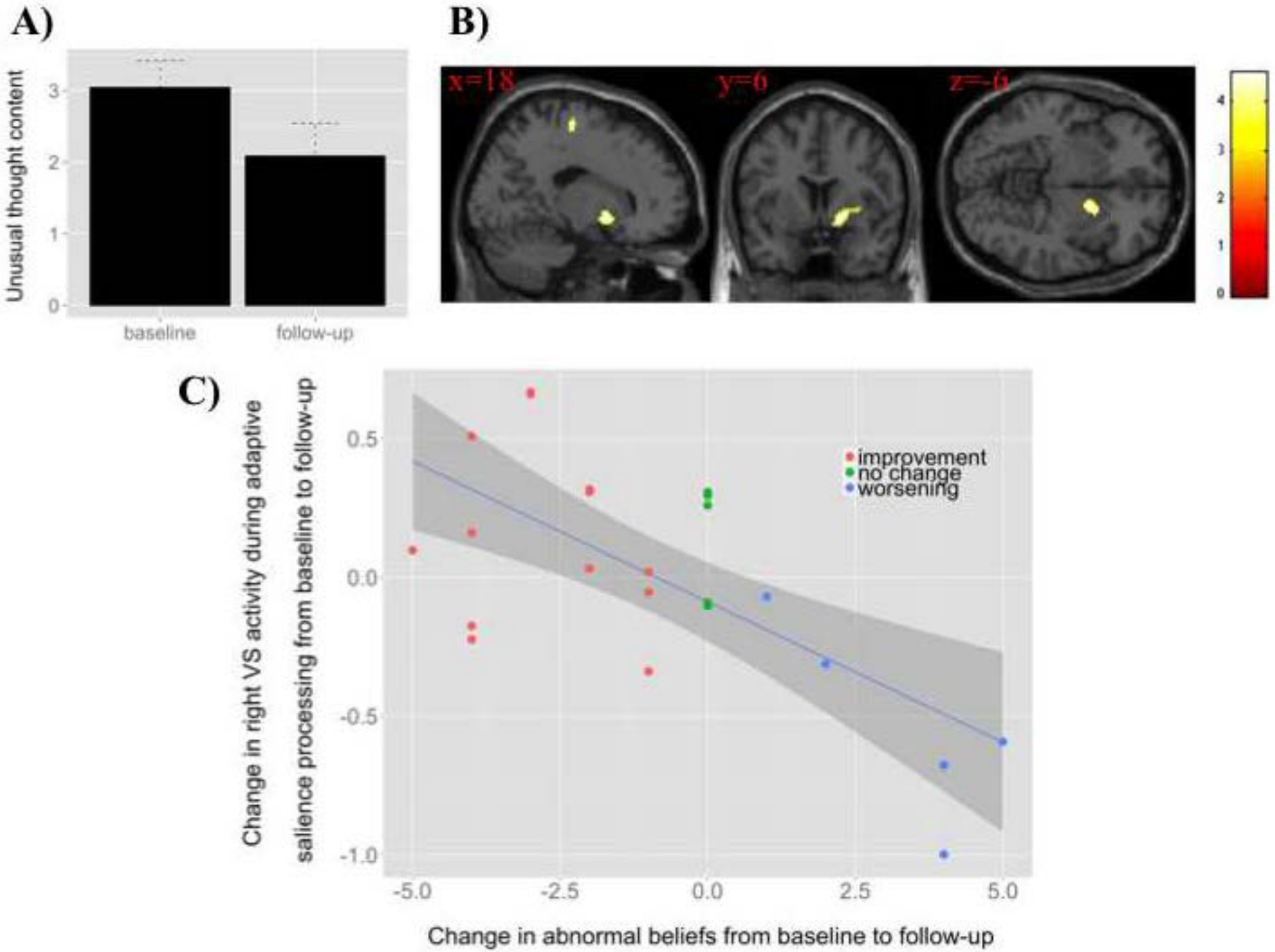


Figure 4



1. Results in healthy controls (n=13) and all subjects at ultra high-risk for psychosis (n=23).

Supplementary Table 1. Group effect across both visits during adaptive reward prediction.

Healthy controls > Ultra high-risk subjects					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.029[a]	/†	12/4/-10	R	3.48
Ventral striatum*	0.038[a]	/†	-14/8/-8	L	3.38
Calcarine sulcus	0.049[b]	322†	12/-74/14	L	4.11
Cuneus			-2/-84/20	L	3.40
Calcarine sulcus			22/-70/6	R	3.38
Midbrain	0.048[b]	311†	-6/-24/-10	L	4.10
Midbrain			6/-22/-10	R	3.61
Middle temporal gyrus	0.047[b]	328†	-48/-54/-4	L	3.77
Superior temporal gyrus	0.7226[b]	53	-46/12/-16	L	3.95
Cuneus	0.1292[b]	223	-14/-74/24	L	3.92
Occipital gyrus			-26/-76/20	L	3.58
Calcarine sulcus	0.2111[b]	175	-8/-78/6	L	3.77
Superior patietal gyrus	0.4378[b]	105	-14/-70/44	L	3.77
Inferior frontal gyrus	0.2600[b]	155	52/10/2	R	3.68
Insula			44/8/6	R	3.65
Cuneus	0.2546[b]	157	18/-96/8	R	3.65
Calcarine sulcus			24/-86/6	R	3.56
Supramarginal gyrus	0.7971[b]	41	54/-36/34	R	3.60

Middle cingulate cortex	0.5099[b]	90	6/6/44	R	3.54
Inferior parietal gyrus	0.5334[b]	35	-30/-44/36	L	3.52
Precentral gyrus	0.6493[b]	65	-50/2/36	L	3.51
Ventral striatum	0.5412[b]	84	12/0/8	R	3.43
Thalamus			12/-14/8	R	3.37
Insula	0.8032[b]	40	-36/-4/12	L	3.41
Postcentral gyrus	0.7103[b]	55	-20/-34/54	L	3.40
Paracentral gyrus			-14/-26/54	L	3.31
Ventral striatum	0.8513[b]	32	-14/8/-8	L	3.38
Thalamus	0.8798[b]	27	-4/-20/12	L	3.34

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. [*]Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level.

No significant effects were found for ultra high-risk subjects > healthy controls.

Supplementary Table 2. Time effect across both groups during adaptive reward prediction.

Follow-up activity > baseline activity					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral part of head of caudate nucleus*	0.0043[a]	/†	-6/10/-2	L	4.09
Ventral part of head of caudate nucleus	0.0911[b]	251	-6/10/-2	L	4.09
Cerebellum	0.8115[b]	39	-4/-36/-46	L	3.84
Thalamus	0.026[b]	383†	22/-18/12	R	3.79
Ventral striatum			14/0/8	R	3.49
Inferior temporal gyrus	0.8833[b]	27	-40/-16/-22	L	3.60
Occipital gyrus	0.7544[b]	48	26/-68/-4	R	3.58
Anterior cingulate cortex	0.6655[b]	62	-18/34/14	L	3.57
Ventral striatum	0.8543[b]	32	-14/-2/6	L	3.56
Inferior frontal gyrus	0.3252[b]	131	46/6/14	R	3.55
Cerebellum	0.8239[b]	37	2/-60/-6	R	3.53
Superior temporal gyrus	0.9104[b]	37	40/-40/16	R	3.52
Lingual gyrus	0.8661[b]	30	-20/-50/2	L	3.44
Insula	0.8115[b]	39	28/34/8	R	3.31

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant effects were found for baseline > follow-up.

Supplementary Table 3. Group effect during adaptive reward prediction at baseline.

Healthy controls > Ultra high-risk subjects					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.025[a]	/†	-16/6/-8	L	3.54
Parahippocampal gyrus	<0.0001[b]	831†	-14/-36/-8	L	4.14
Cerebellum			2/-58/-10	L	3.89
Midbrain			-8/-30/-8	L	3.88
Middle temporal gyrus	0.042[b]	325†	-46/-64/0	L	4.03
Superior temporal gyrus	0.8311[b]	36	-46/14/-14	L	3.90
Precentral gyrus	0.3664[b]	119	-48/-6/40	L	3.76
Insula	0.0735[b]	269	44/6/-2	R	3.64
Ventral striatum	0.7222[b]	53	-16/6/-8	L	3.54
Middle cingulate cortex	0.6773[b]	60	-6/-12/40	L	3.52
Precentral gyrus	0.8556[b]	32	46/-8/44	R	3.50
Supplementary Motor Cortex	0.2877[b]	141	4/6/46	R	3.49
Middle cingulate cortex			10/10/40	R	3.44
Precentral gyrus	0.7868[b]	43	54/-2/36	R	3.48
Ventral striatum	0.8904[b]	26	26/-10/8	R	3.46
Insula	0.5791[b]	76	-40/0/10	L	3.45
Calcarine sulcus	0.8123[b]	39	-12/-80/6	L	3.41
Lingual gyrus			-6/-74/4	L	3.16
Ventral striatum	0.7416[b]	50	-30/-10/-8	L	3.31

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant effects were found for ultra high-risk subjects > healthy controls.

Supplementary Table 4. Group effect during adaptive reward prediction at follow-up.

Healthy controls > Ultra high-risk subjects					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.021[a]	/†	10/18/-2	R	3.60
Ventral striatum*	0.044[a]	/†	-4/14/0	L	3.33
Anterior cingulate cortex	0.0617[b]	293	-16/32/16	L	3.93
Inferior frontal gyrus			-32/20/18	L	3.76
Cerebellum	0.2488[b]	157	-47/-52/-40	L	3.89
Ventral striatum	0.7477[b]	49	10/18/-2	R	3.60
Middle cingulate cortex	0.8531[b]	32	-18/-40/34	L	3.43
Thalamus	0.3540[b]	124	-4/-12/12	L	3.40
Ventral striatum	0.7856[b]	43	-4/14/0	L	3.33
Insula	0.9141[b]	21	-26/-26/30	L	3.30
Cerebellum	0.8351[b]	35	38/-48/-38	R	3.29
Occipital gyrus	0.9141[b]	21	20/-94/10	R	3.26
Occipital gyrus	0.8763[b]	28	-20/-86/14	L	3.24
Insula	0.9192[b]	20	30/-24/28	R	3.21

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected. [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the voxel level. No significant effects were found for ultra high-risk subjects > healthy controls.

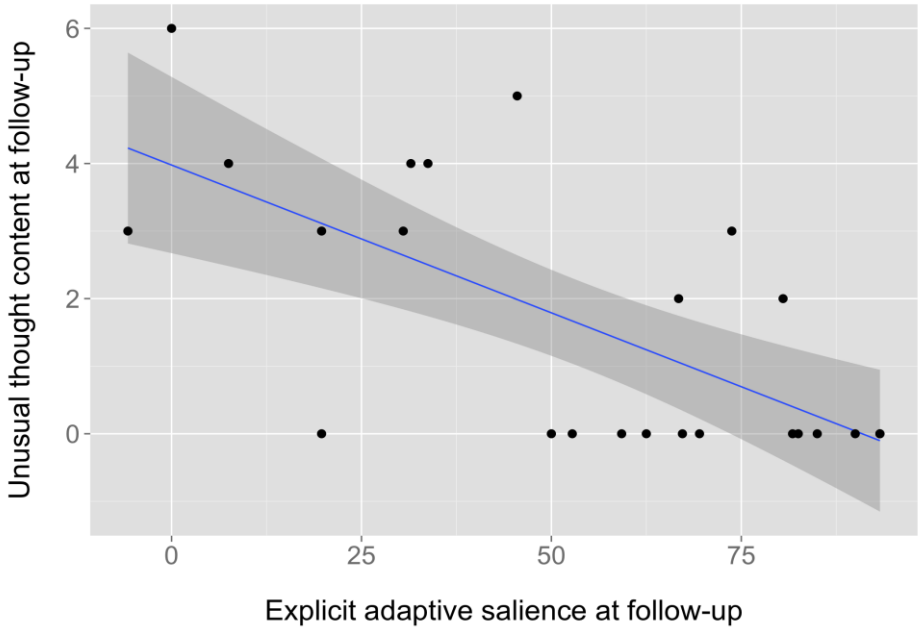
Supplementary Table 5. Negative correlation between longitudinal changes in activation during adaptive reward prediction and change in the severity of abnormal beliefs in ultra high-risk subjects.

Region	P value	Cluster size	MNI co-ordinates (X/Y/Z)	R/L	Z value
Ventral striatum*	0.0171[a]	/†	18/6/-6	R	3.73
Precentral gyrus	0.6126[b]	69	20/-20/64	R	3.79
Ventral striatum	0.2033[b]	160	18/6/-6	R	3.73
Supplementary Motor Cortex	0.026[b]	340†	0/-16/58	R/L	3.65
Supplementary Motor Cortex			-10/-4/70	L	3.52
Supplementary Motor Cortex			4/-6/64	R	3.52

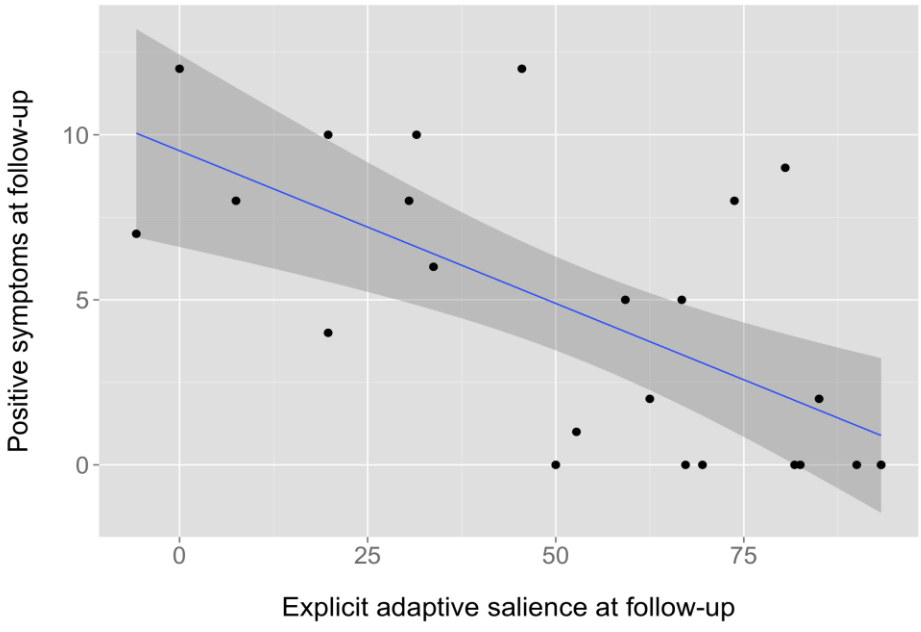
Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected. [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant positive correlations were found.

Supplementary Figure 1. Significant correlations in ultra high-risk subjects between explicit adaptive salience responses (visual analogue scale, VAS) and **(A)** unusual thought content (r=-0.674, p<0.001), **(B)** CAARMS positive symptoms (r=-0.653, p<0.001) and **(C)** global functioning (GAF) (r=0.497, p=0.014) at follow-up.

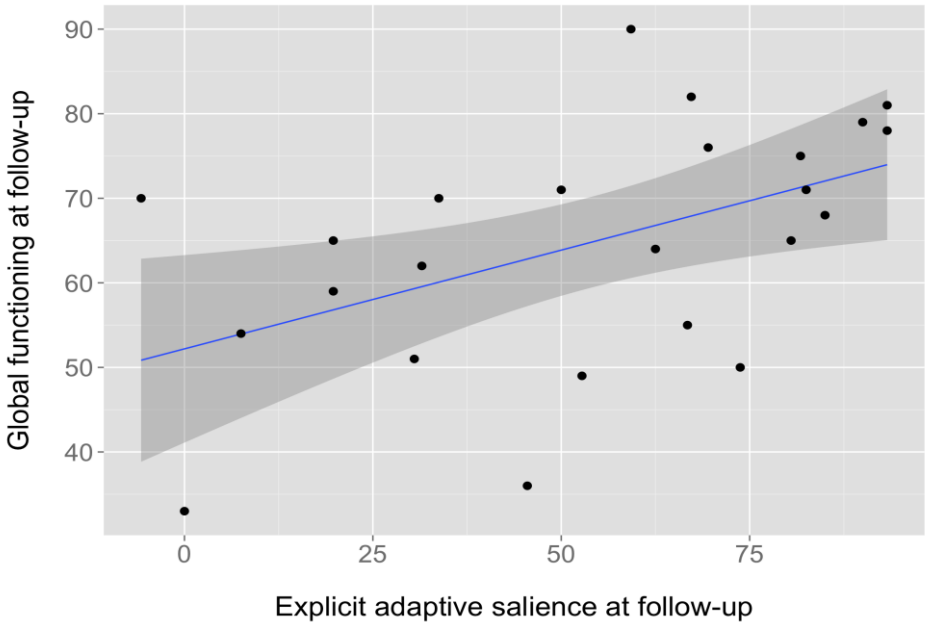
A)



B)



C)


Global functioning at follow-up
90
80
70
60
50
40
0
25
50
75
Explicit adaptive salience at follow-up

2. Results in healthy controls (n=13) and subjects at ultra high-risk for psychosis who did not transit to psychosis (n=22).

A. Behavioural data

Aberrant attribution of salience

Across both visits, UHR subjects showed significantly higher implicit aberrant salience than HCs subjects ($F(1,33)=6.443$, $p=0.016$), and there was a trend for a group x time interaction ($F(1,33)=3.076$, $p=0.089$). There was also a trend for a group x time interaction for explicit aberrant salience ($F(1,33)=3.437$, $p=0.073$). Based on our *a priori* hypotheses we constructed linear contrasts at each time point to test for the predicted group differences in aberrant salience.

At baseline, UHR subjects were more likely than HCs to attribute salience to irrelevant cue features (explicit aberrant salience) ($F(1,33)=5.117$, $p=0.030$), but did not exhibit greater implicit aberrant salience than HCs ($F(1,33)=0.879$, $p=0.335$). At follow-up the group difference in explicit aberrant salience was no longer significant ($F(1,33)=0.073$, $p=0.789$), but HCs had significantly lower implicit aberrant scores than the UHR group ($F(1,33)=11.972$, $p=0.002$) due to a reduction in this measure over time.

Adaptive attribution of salience

Across both visits, the UHR group had lower implicit adaptive salience scores than HCs ($F(1,33)=11.603$, $p=0.002$), as well as lower explicit adaptive salience scores ($F(1,33)=5.763$, $p=0.02$). There was also a strong trend for a group x time interaction for explicit adaptive salience ($F(1,33)=4.086$, $p=0.051$).

At baseline, UHR subjects had significantly lower implicit adaptive salience than HCs ($F(1,33)=14.562$, $p=0.001$) and also exhibited significantly lower explicit adaptive salience ($F(1,33)=9.391$, $p=0.004$). Both of these group differences were no longer significant at follow-up

(implicit adaptive salience: $F(1,33)=3.642$, $p=0.065$; explicit adaptive salience: $F(1,33)=1.446$, $p=0.238$), due to improved scores in the UHR group together with relatively stable performance in HCs.

B. Activation during salience processing

Supplementary Table 6. Group effect across both visits during adaptive reward prediction.

Healthy controls > Ultra high-risk subjects without transition (n=22)					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.025[a]	/†	12/2/8	R	3.54
Ventral striatum*	0.008[a]	/†	-8/8/4	L	3.90
Midbrain	0.024[b]	401†	-6/-24/-10	L	4.25
Midbrain			10/-24/-12	R	3.81
Calcarine sulcus	0.118[b]	231	12/-74/14	R	4.05
Calcarine sulcus			22/-70/6	R	3.31
Superior temporal gyrus	0.619[b]	70	-46/12/-16	L	4.02
Insula	0.514[b]	89	-38/-4/12	L	3.97
Ventral striatum	0.393[b]	115	-8/8/4	L	3.90
Cuneus	0.175[b]	192	-14/-74/24	L	3.87
Occipital gyrus			-26/-76/20	L	3.55
Precentral gyrus	0.253[b]	157	-50/2/36	L	3.80
			-44/-4/36	L	3.59
			-46/-6/50	L	3.25
Insula	0.253[b]	157	44/8/8	R	3.80
Inferior frontal gyrus			52/10/2	R	3.60
Superior patietal gyrus	0.440[b]	104	-14/-70/44	L	3.79
Postcentral gyrus	0.354[b]	125	-20/-34/54	L	3.79
Middle temporal gyrus	0.070[b]	284	-52/-58/-2	L	3.70
Calcarine sulcus	0.272[b]	150	-8/-78/6	L	3.69

Middle cingulate cortex	0.205[b]	177	6/6/44	R	3.68
Cuneus	0.350[b]	126	18/-96/8	R	3.60
Occipital gyrus			26/-86/6	R	3.50
Calcarine sulcus			16/-100/0	R	3.47
Supramarginal gyrus	0.846[b]	33	54/-36/34	R	3.55
Ventral striatum	0.573[b]	78	12/2/8	R	3.54
Inferior parietal gyrus	0.908[b]	22	-30/-44/36	L	3.47
Thalamus	0.840[b]	34	12/-14/8	R	3.46
Supplementary Motor Cortex	0.785[b]	43	2/-16/52	R	3.45
Middle cingulate cortex	0.760[b]	47	-6/-10/40	L	3.41
Middle cingulate cortex			-10/-20/40	L	3.25
Ventral striatum	0.846[b]	33	12/4/-10	R	3.41
Insula	0.816[b]	38	-30/-22/10	L	3.41
Supplementary Motor Cortex	0.918[b]	20	18/-22/52	R	3.37
Ventral striatum	0.840[b]	34	-14/8/-8	L	3.34
Thalamus	0.912[b]	21	-4/-20/12	L	3.34
Precentral gyrus	0.816[b]	38	50/-4/44	R	3.33
Inferior frontal gyrus			56/8/26	R	3.31
Cuneus	0.869[b]	29	-2/-86/20	L	3.32

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. [*]Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level.

No significant effects were found for ultra high-risk subjects > healthy controls.

Supplementary Table 7. Time effect across both groups (healthy controls + ultra high-risk subjects without transition) during adaptive reward prediction.

Follow-up activity > baseline activity					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral part of head of caudate nucleus*	0.006[a]	/†	-4/16/-6	L	3.99
Ventral part of head of caudate nucleus	0.037[a]	/†	19/12/-2	R	3.41
Ventral striatum	0.125[b]	209	-4/-16/-6	L	3.99
Ventral striatum			-18/24/-8	L	3.55
Ventral striatum			-8/12/-12	L	3.43
Inferior frontal gyrus	0.132[b]	211	46/6/14	R	3.85
Insula			38/-16/28	R	3.75
Thalamus	0.0074[b]	513†	22/-18/12	R	3.85
Thalamus			18/-8/10	R	3.69
Ventral striatum			14/0/8	R	3.62
Insula	0.411[b]	108	28/34/8	R	3.72
Inferior frontal gyrus			38/18/10	R	3.19
Lingual gyrus	0.626[b]	68	2/-62/-6	R	3.71
Occipital gyrus	0.683[b]	59	26/-68/-4	R	3.65
Cerebellum	0.868[b]	30	-4/-36/-46	L	3.64
Middle frontal gyrus	0.578	76	-24/34/12	L	3.59
Inferior temporal gyrus	0.902	24	-40/-14/-24	L	3.59
Superior temporal gyrus	0.908[b]	23	40/-40/16	R	3.59
Amygdala	0.923[b]	20	-16/-6/-10	L	3.53
Lingual gyrus	0.923[b]	20	-20/-50/2	L	3.45

Occipital gyrus			-34/-66/2	L	3.27
Ventral striatum	0.908[b]	23	-14/-2/6	L	3.42
Ventral striatum	0.880[b]	28	30/12/6	R	3.24
			24/18/6	R	3.15

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. [*]Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant effects were found for baseline > follow-up.

Supplementary Table 8. Group effect during adaptive reward prediction at baseline.

Healthy controls > Ultra high-risk subjects without conversion (n=22)					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.009[a]	/†	-16/6/-10	L	3.87
Ventral striatum*	0.042[a]	/†	14/2/8	R	3.37
Precentral gyrus	0.069[b]	268	-50/0/38	L	4.37
Precentral gyrus			-42/-10/44	L	3.41
Postcentral gyrus			-44/-20/46	L	3.29
Parahippocampal gyrus	<0.0001[b]	1117†	-12/-36/-8	L	4.32
Cerebellum			0/-40/-12	L	4.16
Midbrain			8/-30/-10	R	3.83
Superior temporal gyrus	0.014[b]	428†	-46/14/-14	L	4.25
Ventral striatum			-16/6/-10	L	3.87
Ventral striatum			-24/-2/-10	L	3.81
Insula	0.106[b]	228	-38/-2/12	L	4.17
Inferior frontal gyrus			-52/6/10	L	3.50
Middle temporal gyrus	0.052[b]	295	-46/-64/0	L	4.08
Middle temporal gyrus			-44/-44/-2	L	3.37
Middle temporal gyrus			-50/-50/-4	L	3.24
Precentral gyrus	0.243[b]	153	46/-8/44	R	4.00
Precentral gyrus			54/2/36	R	3.62
Precentral gyrus			56/8/30	R	3.51

Middle cingulate cortex	0.036[b]	332†	8/10/40	R	3.82
Ventral striatum	0.473[b]	94	26/-10/8	R	3.82
Ventral striatum			22/-2/8	R	3.52
Ventral striatum			12/4/6	R	3.32
Middle cingulate cortex	0.270[b]	144	-6/-12/40	L	3.74
Insula	0.048[b]	304†	44/6/-2	R	3.62
Insula			42/8/6	R	3.61
Inferior frontal gyrus			56/6/8	R	3.55
Precentral gyrus	0.893[b]	26	-22/-28/52	L	3.43
Calcarine sulcus	0.876[b]	29	-12/-80/6	L	3.38
Insula	0.910[b]	23	-32/-22/10	L	3.34

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected, [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant effects were found for ultra high-risk subjects > healthy controls.

Supplementary Table 9. Group effect during adaptive reward prediction at follow-up.

Healthy controls > Ultra high-risk subjects without conversion (n=22)					
Region	**P value**	**Cluster size**	**MNI co-ordinates (X/Y/Z)**	**R/L**	**Z value**
Ventral striatum*	0.025[a]	/†	10/18/-2	R	3.53
Ventral striatum*	0.046[a]	/†	-6/14/0	L	3.31
Anterior cingulate cortex	0.089[b]	255	-16/32/16	L	3.86
Inferior frontal gyrus			-32/20/18	L	3.72
Ventral striatum			-20/14/20	L	3.42
Cerebellum	0.264[b]	151	-46/-52/-40	L	3.82
Ventral striatum	0.811[b]	39	10/18/-2	R	3.53
Middle cingulate cortex	0.914[b]	21	-18/-40/34	L	3.38
Thalamus	0.446[b]	102	-4/-12/12	L	3.35
Thalamus			-12/-4/10	L	3.32
Ventral striatum	0.754[b]	48	-6/14/0	L	3.31
Insula	0.920[b]	20	-26/-26/30	L	3.31
Cerebellum	0.899[b]	24	38/-48/-38	R	3.24

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected. [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the voxel level. No significant effects were found for ultra high-risk subjects > healthy controls.

Supplementary Table 10. Negative correlation between longitudinal changes in activation during adaptive reward prediction and change in the severity of abnormal beliefs in ultra high-risk subjects who did not convert (n=22).

Region	P value	Cluster size	MNI co-ordinates (X/Y/Z)	R/L	Z value
Ventral striatum*	0.021[a]	/†	18/6/-6	R	3.67
Supplementary Motor Cortex	0.001[b]	671†	4/-6/62	R	4.18
Supplementary Motor Cortex			-8/-6/70	L	4.13
Supplementary Motor Cortex			-2/-14/58	L	3.41
Ventral striatum	0.174[b]	171	32/0/2	R	3.75
Ventral striatum			18/6/-6	R	3.67
Ventral striatum	0.745[b]	50	-30/-20/0	L	3.30
Ventral striatum			-26/-8/-4	L	3.18

Results are reported using a cluster-forming threshold $p<0.001$ uncorrected, with an extent threshold of 20 voxels. *Small volume corrected. [a]peak-level FWE-corrected. [b]cluster-level FWE-corrected. † survives FWE correction for multiple comparisons at the cluster or voxel level. No significant positive correlations were found.

Reviewer #1

In this manuscript, the authors explore salience processing deficits in subjects at risk for schizophrenia. This is a hypothesis drawn from the literature published by Kapur, Grace, Heinz, and others, to suggest that alterations in dopamine function in schizophrenia lead to misattribution of salience. The authors have already published extensively on dopamine dysregulation in schizophrenia, and this manuscript goes beyond this to include functional alterations. Interesting, the aberrant salience attribution of the ultra high risk subjects correlated with less activation in the ventral striatum; however, those that normalized over the test period showed increased right ventral striatum activation during salience attribution trials. Overall, the manuscript is well-done and convincing; there are just some minor issues to discuss:

Response: Thank you for the positive feedback and the helpful advice on how to further improve the paper.

1. Whereas the ventral striatum has been associated with reward and affect, it is commonly believed that the associative striatum is involved in salience processing. This is also the striatal region showing greatest amphetamine-induced dopamine release in schizophrenia subjects. Is this a break from this paradigm? Are the authors considering a different conceptualization?

Response: Our conceptualization of salience, as measured by the Salience Attribution Test (SAT), is entirely consistent with the paradigm proposed by Kapur (2003). His theory was cast specifically in terms of motivational salience, which is reliably influenced by dopamine transmission (Berridge and Robinson, 1998) in the ventral striatum. And indeed it was the ventral striatum that showed an aberrant relationship with aberrant salience and positive psychotic symptoms in a previous study from our group ((Roiser et al., 2013), Figure 2).

That said, there is no question that the most reliably identified locus of dopaminergic dysregulation occurs in the associative striatum, as the reviewer correctly highlights. In our view this is an aspect of the neurobiology of aberrant salience that remains to be clarified. We have now added material to the discussion on this point:

Pages 15 and 16:

"Taken with longitudinal PET imaging findings that changes in dopamine synthesis capacity in dorsal (associative) striatum are associated with change in clinical state (Howes et al., 2011a), our findings suggest that alterations in both ventral and dorsal striatum are linked to symptom change. A possible mechanism could be that hyperactive inputs from the hippocampus to the ventral striatum in psychosis may impact dopaminergic neurons that project to more dorsal (associative) striatal areas and thereby affect dorsal striatum related salience processing (Haber, 2003, Lodge and Grace, 2011, 2012, Modinos et al., 2014)."

Page 17:

"Furthermore, in accordance with aberrant salience model (Kapur, 2003), the SAT has been designed to measure abnormal motivational (reward) salience processing in psychosis and its relation to dopamine dysregulation in the VS. However, motivation is not the only form of salience (Winton-Brown et al., 2014), and it would be important to test ventral and dorsal (associative) striatal activation in psychosis during other forms of salience processing that are not measured using speeded response tasks."

2. Is there a concern that the paradigm involved monetary rewards and speed of response, which would favor a ventral striatal reward-related activation more than a salience attribution activation? Was there a regional difference in activation based on reward versus salience?

Response: The SAT was designed to parallel animal research in which response vigor is the method used to index motivational salience processing (during Pavlovian to instrumental transfer). As discussed in our response to the reviewer's previous response, in Kapur's theory, which was the basis for the development of the SAT, salience is cast in terms of motivational processing, i.e. reward. While we agree that motivation is not the only form of salience, using the SAT we cannot identify other forms of salience. We have now added material to the discussion on this point:

Page 17:

"Furthermore, in accordance with aberrant salience model (Kapur, 2003), the SAT has been designed to measure abnormal motivational (reward) salience processing in psychosis and its relation to dopamine dysregulation in the VS. However, motivation is not the only form of salience (Winton-Brown et al., 2014), and it would be important to test ventral and dorsal (associative) striatal activation in psychosis during other forms of salience processing that are not measured using speeded response tasks."

We also changed the title to:

"Longitudinal alterations in motivational salience processing in ultra high-risk subjects for psychosis"

3.The model regarding ventral striatal activation altering midbrain-dorsal striatal dopamine activation is intriguing, and is in concert with basic neuroscience investigations, as outlined in recent reviews based on animal models. It is also consistent with a preferential projection of the hippocampus (hyperactive in psychosis) to ventral striatal regions rather than dorsal striatal regions. A reference to this interaction would be useful, and consistent with the Haber loop model alluded to in this paper.

Response: Based on your suggestion, we adjusted this statement on pages 15 and 16 as follows:

"A possible mechanism could be that hyperactive inputs from the hippocampus to the ventral striatum in psychosis may impact dopaminergic neurons that project to more dorsal (associative) striatal areas and thereby affect dorsal striatum related salience processing (Haber, 2003, Lodge and Grace, 2011, 2012, Modinos et al., 2014)."

Reviewer #2

This is a well written, succinct work on alterations in salience processing in UHR subjects at baseline and after a period of follow-up. With some revision, this paper will add a substantial contribution to the current literature.

While the primary purpose seems to be focussed on aberrant salience, the lack of results supporting aberrant salience processing (i.e. irrelevant attribution) is striking in this study. Although UHR group showed lower explicit aberrant salience processing at presentation (which improved after follow up), there was no correlation between aberrant salience and symptoms, no differences in brain activation during aberrant reward prediction, no relationship between changes in clinical features and brain activation during aberrant reward prediction. Whereas they do have some results for adaptive salience processing. i.e. negative correlation between severity of abnormal beliefs as well as positive symptoms and adaptive salience, positive correlation between GAF scores and adaptive salience, and lowered activation in the ventral striatum during adaptive reward prediction at presentation and follow up compared to HC, correlation between degree of improvement in abnormal beliefs and increase in activation in the ventral striatum during adaptive reward prediction. I find this quite intriguing.

Response: Thank you for your thorough and helpful assessment. In what follows, we have addressed the issues you raised and have incorporated the changes you proposed.

1. I think these results are not fully in line with the original aberrant salience model. These results seem to indicate that a UHR state is characterized by reduced attribution of salience to relevant stimuli (and accompanying lowered activation in the VS) as opposed to increased attribution of salience to irrelevant stimuli. This is interesting, and in this regard, the results speak more closely to the salience network based models of attentional processing and network switching when dealing with a relevant task at hand. I think the authors should perhaps comment more on this in the discussion section (they do briefly touch upon this but do not include the most relevant arguments and works in this regard).

Response: We like to mention that the UHR subjects in this study were indeed more likely to attribute salience to irrelevant stimuli than HCs at clinical presentation (see behavioural results), which is consistent with the aberrant salience model. Our results thus replicate a previous report of increased explicit aberrant salience in an independent UHR sample (Roiser et al., 2013). However, you're right that we found no group differences in brain activation during aberrant reward prediction. There are different potential explanations for this lack of effect and one is indeed the low conversion rate in our UHR sample as you have mentioned in the point below. We summarized these possibilities on pages 16 and 17:

"Some limitations of our study merit comment. The sample sizes were modest, largely because inclusion required that participants completed multi-modal neuroimaging assessments at both baseline and follow-up. The modest group sizes may thus have accounted for the absence of significant group differences in activation during aberrant salience processing. A further consideration is that at the time of writing, only one UHR subject had developed a psychotic disorder (all results remained after excluding this subject, see supplementary information 2A and B for details), precluding any examination of the relationship between abnormal salience processing and the risk of transition to psychosis. In this regard, it is possible that the low conversion rate in our UHR sample may explain the lack of alterations in brain activation during aberrant salience processing. Future large-scale studies with a meaningful ratio between converters and non-converters are required to test if functional brain alterations during aberrant reward prediction are evident in UHR subjects who later develop psychosis or if the risk of transition to psychosis is more related to impaired activation when dealing with a relevant task at hand (i.e. adaptive reward prediction). Furthermore, in accordance with aberrant salience model (Kapur, 2003), the SAT has been designed to measure abnormal motivational (reward) salience processing in psychosis and its relation to dopamine dysregulation in the VS. However, motivation is not the only form of salience (Winton-Brown et al., 2014), and it would be important to test ventral and dorsal (associative)

striatal activation in psychosis during other forms of salience processing that are not measured using speeded response tasks."

Nevertheless, we agree that the lack of significant brain differences during aberrant salience processing is intriguing and may indicate that psychosis is more closely associated with deficits in processing of contextually relevant information. According to your suggestion, we included a statement on page 16 in the revised manuscript:

"However, the SAT is a complex task that also involves sustained attention, maintaining stimulus information in memory, decision-making and response selection (Roiser et al., 2013), and the UHR state is associated with a broad range of cognitive impairments (Fusar-Poli et al., 2012). We therefore speculate that this finding in the supplementary motor cortex may be related to alterations in one or more of these processes, possibly secondary to changes in striatal function. Furthermore, UHR subjects also showed reduced activation in the calcarine sulcus, cuneus, midbrain and middle temporal gyrus across both visits during the attribution of salience to relevant stimuli, as well as reduced activation in the parahippocampal gyrus, cerebellum, midbrain, middle temporal gyrus, middle and anterior cingulate cortex, inferior frontal gyrus and insula at baseline and/or follow-up (see supplementary information 1 and 2B for more details). Together with the striatum, integration of these regions is important to sustain emotion and cognition, especially during the detection and processing of salient information (Menon and Uddin, 2010, Seeley et al., 2007). Dysfunction of this network and abnormal network switching when dealing with a relevant task at hand has been proposed to underlie the formation of psychotic symptoms (Palaniyappan and Liddle, 2012, Palaniyappan et al., 2013, Schmidt et al., 2016)."

And we also adjusted the limitation section on page 17:

"In this regard, it is possible that the low conversion rate in our UHR sample may explain the lack of alterations in brain activation during aberrant salience processing. Future large-scale studies with a meaningful ratio between converters and non-converters are required to test if functional brain alterations during aberrant reward prediction are evident in UHR subjects who later develop psychosis or if the risk of transition to psychosis is more related to impaired activation when dealing with a relevant task at hand (i.e. adaptive reward prediction).

2. Given the fact that behavioural scores for salience attribution (both relevant and irrelavant) improved at follow up, and positive and negative symptoms improved significantly at follow-up, and given the lack of relationship between aberrant salience and symptoms or neural activity, is it possible that this particular UHR group is less likely to transition to psychosis? Could this perhaps explain lack of alterations in aberrant salience processing? Would the findings perhaps be different in subjects that would go on to experience a full blown psychotic episode? It would be interesting to investigate this in a large scale longitudinal study.

Response: It conceivable that the low conversion rate of our UHR sample (only one subject transited to psychosis over the follow-up period) may explain the lack of alterations in brain activation during aberrant reward prediction. We added a limitation to this issue on pages 16 and 17 (please see our response to your point above).

"Some limitations of our study merit comment. The sample sizes were modest, largely because inclusion required that participants completed multi-modal neuroimaging assessments at both baseline and follow-up. The modest group sizes may thus have accounted for the absence of significant group differences in activation during aberrant salience processing. A further consideration is that at the time of writing, only one UHR subject had developed a psychotic disorder (all results remained after excluding this subject, see supplementary information 2A and B for details), precluding any examination of the relationship between abnormal salience processing and the risk of transition to psychosis. In this regard, it is possible that the low conversion rate in our UHR sample may explain the lack of alterations in brain activation during aberrant salience processing. Future large-scale studies with a meaningful ratio between converters and non-

converters are required to test if functional brain alterations during aberrant reward prediction are evident in UHR subjects who later develop psychosis or if the risk of transition to psychosis is more related to impaired activation when dealing with a relevant task at hand (i.e. adaptive reward prediction).”

3. The introduction could be more representative of the general literature, rather than focussing mostly on the authors prior work.

Response: According to your suggestion we have expanded the introduction in the revised manuscript (within the admissible word range).

4. In the results for implicit aberrant salience, there was no difference in groups at baseline but at follow up HCs show reduced implict aberrant salience and this is simply explained as a 'reduction in this measure over time'. Why does this measure reduce over time in HCs?

Response: A potential suggestion would be that HCs improve on the task over time due to a practice effect by which HCs become better at attending to the irrelevant stimulus dimension. However, this is the first report of such an effect in HCs and it requires replication in future studies.

5. It seems that the authors have used scrubbing/artrepair. Any difference in % of interpolated volumes between groups?

Response: We did not use scrubbing/artrepair. In total, we interpolated 24 volumes (5.1%) in the HC group and 7 volumes (1.48%) in the UHR group across both runs.

We added this information on page 7:

“Additionally, all images were checked for movement artefacts, and all scans with more than 5 mm deviation from the previous scan in any dimension, resulting in corrupted volumes, were excluded and replaced with the average of the neighbouring volumes (5.1% in HCs and 1.5% in UHRs).”

6. I feel like the last sentence in the discussion section is a bit misleading.. only adaptive salience was associated with VS function and symptom improvement and although the degree of improvement in abnormal beliefs over time was associated with the longitudinal increase in activation during adaptive reward prediction in the right VS, UHR still showed significantly lower VS activation compared to HC even at follow-up, so it may not be appropriate to say 'normalisation of VS function'.

Response: We agree and reformulated the conclusion at the end of the discussion as follows:

“In summary, this study shows that UHR subjects exhibit behavioural deficits in both adaptive and aberrant salience processing at clinical presentation, which disappeared along with the remission of attenuated psychotic symptoms over the follow-up period. Our results further indicate ventral striatal hypoactivation in UHR subjects during adaptive reward prediction at baseline and follow-up and that the amelioration of abnormal beliefs over the follow-up period is linked to a longitudinal increase in VS activation during adaptive reward prediction.”

Reviewer #3

The topic of the study is interesting and worth investigating, study uses a well-established SAT task, methodology and the results are described in enough detail. Limitations are adequately addressed. In my opinion study the paper should be published, two minor points for consideration:

Response: Thank you for this encouraging comment and the helpful advice on how to further improve the paper.

1. only the findings regarding ventral striatum and supplementary motor cortex activations are discussed, while the between-group differences in activation of other structures during the adaptive reward prediction are reported as well (also at the FWE corrected threshold) with no functional significance of the findings being discussed.

Response: We agree that there are many other regions showing significant (FWE-corrected) between-group differences during adaptive reward prediction. In particular, we also found significant group effects across both visits in the calcarine sulcus, cuneus, midbrain and middle temporal gyrus, as well as group effects at baseline and follow-up in the parahippocampal gyrus, cerebellum, midbrain, middle temporal gyrus, middle and anterior cingulate cortex, inferior frontal gyrus and insula (see supplementary information 1 and 2B for more detail). According to your suggestion, we discuss these findings on page 16 in the revised manuscript:

"However, the SAT is a complex task that also involves sustained attention, maintaining stimulus information in memory, decision-making and response selection (Roiser et al., 2013), and the UHR state is associated with a broad range of cognitive impairments (Fusar-Poli et al., 2012). We therefore speculate that this finding in the supplementary motor cortex may be related to alterations in one or more of these processes, possibly secondary to changes in striatal function. Furthermore, UHR subjects also showed reduced activation in the calcarine sulcus, cuneus, midbrain and middle temporal gyrus across both visits during the attribution of salience to relevant stimuli, as well as reduced activation in the parahippocampal gyrus, cerebellum, midbrain, middle temporal gyrus, middle and anterior cingulate cortex, inferior frontal gyrus and insula at baseline and/or follow-up (see supplementary information 1 and 2B for more details). Together with the striatum, integration of these regions is important to sustain emotion and cognition, especially during the detection and processing of salient information (Menon and Uddin, 2010, Seeley et al., 2007). Dysfunction of this network and abnormal network switching when dealing with a relevant task at hand has been proposed to underlie the formation of psychotic symptoms (Palaniyappan and Liddle, 2012, Palaniyappan et al., 2013, Schmidt et al., 2016)."

2. no information regarding the UHR inclusion criteria (APS/BLIPS/APS+GRD) for the final sample included in the analyses is given.

Response: Thank you for raising this point. We added this information on page 7:

"In the final sample of 23 UHR subjects, 19 subjects were included based on BLIPS, three based on BLIPS and one based on APS+GRD."

3. Moreover, have the findings been re-examined after the exclusion of the participant who transitioned to psychosis?

Response: Based on your suggestion we repeated the analyses after excluding this subject. These results can be found in the supplementary information 2. We also adjusted the result section and the statement on pages 16 and 17:

"A further consideration is that at the time of writing, only one UHR subject had developed a psychotic disorder (all results remained after excluding this subject, see supplementary information

2A and B for details), ***precluding any examination of the relationship between abnormal salience processing and the risk of transition to psychosis."***

References

Berridge, K. C. & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* **28**, 309-69.
Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat* **26**, 317-30.
Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* **160**, 13-23.
Lodge, D. J. & Grace, A. A. (2011). Hippocampal dysregulation of dopamine system function and the pathophysiology of schizophrenia. *Trends Pharmacol Sci* **32**, 507-13.
Lodge, D. J. & Grace, A. A. (2012). Divergent activation of ventromedial and ventrolateral dopamine systems in animal models of amphetamine sensitization and schizophrenia. *Int J Neuropsychopharmacol* **15**, 69-76.
Menon, V. & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* **214**, 655-67.
Modinos, G., Allen, P., Grace, A. A. & McGuire, P. (2014). Translating the MAM model of psychosis to humans. *Trends Neurosci*.
Palaniyappan, L. & Liddle, P. F. (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J Psychiatry Neurosci* **37**, 17-27.
Palaniyappan, L., Simmonite, M., White, T. P., Liddle, E. B. & Liddle, P. F. (2013). Neural primacy of the salience processing system in schizophrenia. *Neuron* **79**, 814-28.
Roiser, J. P., Howes, O. D., Chaddock, C. A., Joyce, E. M. & McGuire, P. (2013). Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophr Bull* **39**, 1328-36.
Schmidt, A., Palaniyappan, L., Smieskova, R., Simon, A., Riecher-Rössler, A., Lang, U. E., Fusar-Poli, P., McGuire, P. & Borgwardt, S. J. (2016). Dysfunctional insular connectivity during reward prediction in patients with first-episode psychosis. *J Psychiatry Neurosci* **41**, 150234.
Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L. & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci* **27**, 2349-56.

Institute of Psychiatry

at The Maudsley

KING'S College LONDON Founded 1829

University of London

Prof. Robin M. Murray
Editor-in Chief (UK), Psychological Medicine
Institute of Psychiatry, Psychology & Neuroscience
London, UK

London, August 2016

Dear Professor Murray,

We have carefully considered the reviewers' comments on our submitted manuscript (PSM-D-16-00500: Longitudinal alterations in salience processing in ultra high-risk subjects for psychosis) and have now completed the revision. Our responses to the critiques are provided in a point-by-point fashion. For your convenience, all changes made to the original manuscript are marked in red color in the revised manuscript. We believe that the manuscript is now considerably improved and has adequately addressed the concerns raised by the reviewers.

We hope that that the manuscript now finds the reviewers' and your approval and look forward to hearing your final decision.

Sincerely yours,
André Schmidt